



THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

### The CALYMHA survey: Ly escape fraction and its dependence on galaxy properties at $z = 2.23$

**Citation for published version:**

Matthee, J, Sobral, D, Oteo, I, Best, P, Smail, I, Röttgering, H & Paulino-Afonso, A 2016, 'The CALYMHA survey: Ly escape fraction and its dependence on galaxy properties at  $z = 2.23$ ', *Monthly Notices of the Royal Astronomical Society*, vol. 458, no. 1, pp. 449-467. <https://doi.org/10.1093/mnras/stw322>

**Digital Object Identifier (DOI):**

[10.1093/mnras/stw322](https://doi.org/10.1093/mnras/stw322)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Publisher's PDF, also known as Version of record

**Published In:**

Monthly Notices of the Royal Astronomical Society

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.





# The CALYMHA survey: Ly $\alpha$ escape fraction and its dependence on galaxy properties at $z = 2.23$

Jorjy Matthee,<sup>1★</sup> David Sobral,<sup>1,2,3</sup> Iván Oteo,<sup>4,5</sup> Philip Best,<sup>4</sup> Ian Smail,<sup>6</sup> Huub Röttgering<sup>1</sup> and Ana Paulino-Afonso<sup>2</sup>

<sup>1</sup>*Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands*

<sup>2</sup>*Instituto de Astrofísica e Ciências do Espaço, Universidade de Lisboa, OAL, Tapada da Ajuda, P-1349-018 Lisboa, Portugal*

<sup>3</sup>*Department of Physics, Lancaster University, Lancaster LA1 4YB, UK*

<sup>4</sup>*Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK*

<sup>5</sup>*European Southern Observatory, Karl-Schwarzschild-Str. 2, D-85748 Garching, Germany*

<sup>6</sup>*Centre for Extragalactic Astronomy, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK*

Accepted 2016 February 8. Received 2016 February 8; in original form 2015 December 23

## ABSTRACT

We present the first results from our CALibrating LYman  $\alpha$  with H $\alpha$  (CALYMHA) pilot survey at the Isaac Newton Telescope. We measure Ly $\alpha$  emission for 488 H $\alpha$  selected galaxies at  $z = 2.23$  from High- $z$  Emission Line Survey in the COSMOS and UDS fields with a specially designed narrow-band filter ( $\lambda_c = 3918 \text{ \AA}$ ,  $\Delta\lambda = 52 \text{ \AA}$ ). We find 17 dual H $\alpha$ -Ly $\alpha$  emitters [ $f_{\text{Ly}\alpha} > 5 \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$ , of which five are X-ray active galactic nuclei (AGN)]. For star-forming galaxies, we find a range of Ly $\alpha$  escape fractions ( $f_{\text{esc}}$ , measured with 3 arcsec apertures) from 2 to 30 per cent. These galaxies have masses from  $3 \times 10^8 M_\odot$  to  $10^{11} M_\odot$  and dust attenuations  $E(B - V) = 0\text{--}0.5$ . Using stacking, we measure a median escape fraction of  $1.6 \pm 0.5$  per cent ( $4.0 \pm 1.0$  per cent without correcting H $\alpha$  for dust), but show that this depends on galaxy properties. The stacked  $f_{\text{esc}}$  tends to decrease with increasing star formation rate and dust attenuation. However, at the highest masses and dust attenuations, we detect individual galaxies with  $f_{\text{esc}}$  much higher than the typical values from stacking, indicating significant scatter in the values of  $f_{\text{esc}}$ . Relations between  $f_{\text{esc}}$  and UV slope are bimodal, with high  $f_{\text{esc}}$  for either the bluest or reddest galaxies. We speculate that this bimodality and large scatter in the values of  $f_{\text{esc}}$  is due to additional physical mechanisms such as outflows facilitating  $f_{\text{esc}}$  for dusty/massive systems. Ly $\alpha$  is significantly more extended than H $\alpha$  and the UV  $f_{\text{esc}}$  continues to increase up to at least 20 kpc ( $3\sigma$ , 40 kpc [ $2\sigma$ ]) for typical star-forming galaxies and thus the aperture is the most important predictor of  $f_{\text{esc}}$ .

**Key words:** galaxies: evolution – galaxies: high-redshift – galaxies: ISM.

## 1 INTRODUCTION

The Lyman  $\alpha$  (Ly $\alpha$ ) emission line (rest-frame  $1216 \text{ \AA}$ ) has emerged as a powerful tool to study distant galaxies, since it is intrinsically the brightest emission line in HII regions and redshifted into optical wavelengths at  $z > 2$ . As a result, the Ly $\alpha$  line has been used to spectroscopically confirm the highest redshift galaxies (Oesch et al. 2015; Zitrin et al. 2015), select samples of galaxies with narrow-band (NB) filters (e.g. Ouchi et al. 2008; Matthee et al. 2015), find extremely young galaxies (e.g. Kashikawa et al. 2012; Sobral et al. 2015b), study the interstellar, circumgalactic and intergalactic medium (e.g. Rottgering et al. 1995; Cantalupo et al. 2014;

Swinbank et al. 2015) and probe the epoch of reionization (e.g. Ouchi et al. 2010; Dijkstra 2014).

However, due to the resonant nature of Ly $\alpha$ , it is unknown what the observed strength of the Ly $\alpha$  emission line actually traces. While Ly $\alpha$  photons are emitted as recombination radiation in HII regions, where ionizing photons originate from star formation or AGN activity, Ly $\alpha$  photons can also be emitted by collisional ionization due to cooling (e.g. Rosdahl & Blaizot 2012) and shocks. Most importantly, only a small amount of neutral hydrogen is needed to get an optical depth of 1 (with column densities of  $\sim 10^{14} \text{ cm}^{-2}$ ; Hayes 2015). Therefore, Ly $\alpha$  photons are likely to undergo numerous scattering events before escaping a galaxy. This increases the likelihood of Ly $\alpha$  being absorbed by dust and also leads to a lower surface brightness (SB; detectable as Ly $\alpha$  haloes; e.g. Steidel et al. 2011; Momose et al. 2014; Wisotzki et al. 2015) and diffusion in

\* E-mail: matthee@strw.leidenuniv.nl

wavelength space, altering line profiles (e.g. Verhamme et al. 2008; Dijkstra 2014).

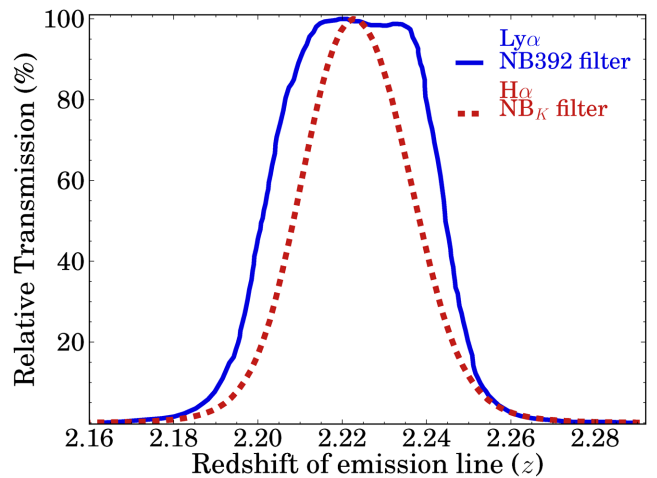
In order to use Ly $\alpha$  to search for and study galaxies in the early Universe, it is of key importance to directly measure the fraction of intrinsically produced Ly $\alpha$  (the Ly $\alpha$  escape fraction,  $f_{\text{esc}}$ ), and to understand how that may depend on galaxy properties. Under the assumption of case B recombination radiation,  $f_{\text{esc}}$  can be measured by comparing the Ly $\alpha$  flux with H $\alpha$ . H $\alpha$  (rest-frame 6563 Å) is not a resonant line and typically only mildly affected by dust in a well-understood way (e.g. Garn & Best 2010). Measurements of both Ly $\alpha$  and H $\alpha$  can thus improve the understanding of what Ly $\alpha$  actually traces by comparing  $f_{\text{esc}}$  with other observables as mass, dust content, kinematics, or Ly $\alpha$  line properties [such as the Equivalent Width (EW) and profile].

It is in principle possible to estimate the intrinsic Ly $\alpha$  production using other tracers of the ionizing photon production rate [i.e. other star formation rate (SFR) indicators]. However, these all come with their own uncertainties and assumptions. For example, studies using H $\beta$  (e.g. Ciardullo et al. 2014) and UV-selected samples (e.g. Gronwall et al. 2007; Nilsson et al. 2009; Blanc et al. 2011; Cassata et al. 2015) suffer from more significant and uncertain dust corrections, and may select a population which tends to be less dusty (e.g. Oteo et al. 2015). UV-based studies are furthermore dependent on uncertainties regarding SED modelling, and on assumptions on the time-scales (UV typically traces SFR activity over a 10 times longer time-scale than nebular emission lines; e.g. Boquien, Buat & Perret 2014). Estimates using the far-infrared (Wardlow et al. 2014; Kusakabe et al. 2015) suffer from even larger assumptions on the time-scales. Remarkably though, most studies find a consistent value of  $f_{\text{esc}} \sim 30$  per cent for Ly $\alpha$  emitters (LAEs), and lower for UV-selected galaxies,  $\sim 3$ –5 per cent (e.g. Hayes et al. 2011).

Locally, it has been found that the Ly $\alpha$  escape fraction anti-correlates with dust attenuation (Cowie, Barger & Hu 2010; Atek et al. 2014), although the large scatter indicates that there are other regulators of Ly $\alpha$  escape, such as outflows (e.g. Kunth et al. 1998; Atek et al. 2008; Rivera-Thorsen et al. 2015). However, these locally studied galaxies have been selected in different ways than typical high-redshift galaxies. Green pea galaxies (selected by their strong nebular [O III] emission) have recently been studied as local analogues for high-redshift LAEs (e.g. Henry et al. 2015; Yang et al. 2015). These studies find indications that the escape fraction correlates with H I column density, and that is also related to galactic outflows and dust attenuation. However, the sample sizes and the dynamic range are still significantly limited.

At higher redshift, it is challenging to measure the Ly $\alpha$  escape fraction, as H $\alpha$  can only be observed up to  $z \sim 2.8$  from the ground, while Ly $\alpha$  is hard to observe at  $z < 2$ . Therefore,  $z \sim 2.5$  is basically the only redshift window where we can directly measure both Ly $\alpha$  and H $\alpha$  with current instrumentation. This experiment has been performed by Hayes et al. (2010), who found a global average escape fraction of  $5 \pm 4$  per cent. The escape fraction is obtained by comparing integrated H $\alpha$  and Ly $\alpha$  luminosity functions (see also Hayes et al. 2011), so the results depend on assumptions on the shape of the luminosity function, integration limits, etc. Recently, Oteo et al. (2015) found that only 4.5 per cent of the H $\alpha$  emitters (HAEs) covered by Nilsson et al. (2009) are detected as LAEs, indicating a similar escape fraction.

In order to increase the sample size and study dependencies on galaxy properties, we have recently completed the first phase of our CALYMA survey: CALibrating LYMan  $\alpha$  with H $\alpha$ . This survey combines the  $z = 2.23$  HAEs from High- $z$  Emission Line Survey (HiZELS) (Sobral et al. 2013) with Ly $\alpha$  measurements using a



**Figure 1.** Filter transmission curves for the NBs used to measure H $\alpha$  (NB $_K$ ) and Ly $\alpha$  (NB392). The NB392 filter is designed to provide complete coverage of the redshifts at which HAEs can be selected in NB $_K$ , from  $z = 2.20$  to 2.25. The Ly $\alpha$  emission for all our HAEs is covered even if it is shifted by  $\pm 600$  km s $^{-1}$ . Depending on the specific redshift, the filter transmission varies between the two lines, such that Ly $\alpha$  is typically over-estimated with respect to H $\alpha$  (see Section 3.4.2). We statistically correct for this in stacked or median measurements.

custom-made NB filter (see Fig. 1). The observations from our pilot survey presented here cover the full COSMOS field and a major part of the UDS field, and are described in Sobral et al. (in preparation). The aim of this paper is to measure the escape fraction for the H $\alpha$  selected sources, and measure median stacked escape fractions in multiple subsets in order to understand which galaxy properties influence  $f_{\text{esc}}$ .

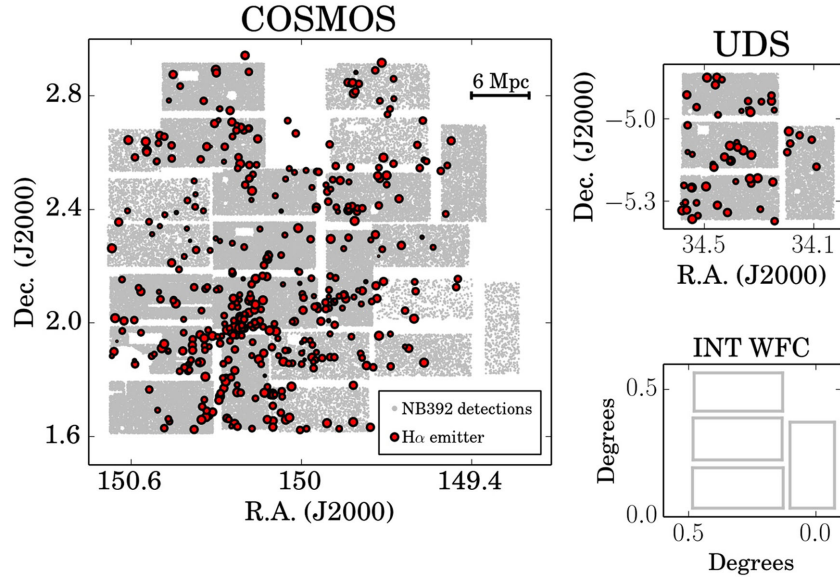
The structure of this paper is as follows. In Section 2 we present the sample of  $z = 2.23$  HAEs and the Ly $\alpha$  observations. We describe our method to measure Ly $\alpha$  line-flux and escape fraction and galaxy properties in Section 3, while Section 4 describes our stacking method. Section 5 presents the Ly $\alpha$  properties of individual galaxies. We explore correlations between  $f_{\text{esc}}$  and galaxy properties in Section 6 and study extended Ly $\alpha$  emission in Section 7. Our results are compared with other studies in Section 8 and we summarize our results and present our conclusions in Section 9.

Throughout the paper, we use a  $\Lambda$  cold dark matter cosmology with  $H_0 = 70$  km s $^{-1}$  Mpc $^{-1}$ ,  $\Omega_M = 0.3$  and  $\Omega_\Lambda = 0.7$ . Magnitudes are given in the AB system and measured in 3 arcsec diameter apertures, unless noted otherwise. At  $z = 2.23$ , 1 arcsec corresponds to a physical scale of 8.2 kpc. We use a Chabrier (2003) IMF to obtain stellar masses and star formation rates.

## 2 SAMPLE AND OBSERVATIONS

### 2.1 Sample of H $\alpha$ emitters

We use a sample of HAEs at  $z = 2.23$  in the COSMOS and UDS fields selected from the HiZELS (Geach et al. 2008; Sobral et al. 2009; Best et al. 2013; Sobral et al. 2013) using NB imaging in the  $K$  band with the United Kingdom Infra Red Telescope (UKIRT). HAEs are identified using  $BzK$  and  $BRU$  colours and photometric redshifts, as described in Sobral et al. (2013). These HAEs are selected to have  $EW_{0, H\alpha + [N III]} > 25$  Å. In total, there are 588 HAEs at  $z = 2.23$  in COSMOS, of which 552 are covered by our Ly $\alpha$  survey area. We remove 119 HAEs because they are found in noisy regions of the Ly $\alpha$  coverage, resulting in a sample of 433 HAEs in



**Figure 2.** Positions on the sky in COSMOS and UDS of the HAEs from Sobral et al. (2013) in red points, where the size of the symbols scales with observed  $H\alpha$  luminosity. Our  $\sim 2 \text{ deg}^2$  coverage includes a wide range of environments, with number density of sources on the sky varying over orders of magnitudes, overcoming cosmic variance (see e.g. Sobral et al. 2015a). The grey points show all detections in our NB392 observations, after conservative masking of noisy regions due to the dithering pattern. It can be seen that some pointings are shallower with a lower number density of sources, and that we masked regions around bright stars and severe damages to one of the chips. After our conservative masking, we use a total area of  $1.208 \text{ deg}^2$  in COSMOS and  $0.224 \text{ deg}^2$  in UDS. We also show the four detector chips of the WFC on the INT with a total field of view of  $\sim 0.25 \text{ deg}^2$ .

COSMOS. The UDS sample consists of 184 HAEs, of which 55 are observed to sufficient signal-to-noise ratio in the Isaac Newton Telescope (INT) imaging (local background of  $23.5, 3\sigma$ , or deeper). This means that our total sample includes 488 HAEs, shown in Fig. 2.

The multi-wavelength properties of the HAEs are discussed in Oteo et al. (2015), showing that the  $H\alpha$  selection incorporates the full diversity of star-forming galaxies (SFGs; e.g. in their figs 5 and 6), while selections based on the Lyman break or the  $Ly\alpha$  emission line miss significant parts of the SFG population at  $z = 2.23$ . Furthermore, although our sample of galaxies contains strongly star-bursting systems, the majority is not biased towards these rare sources. Our sample is dominated by typical galaxies which are on the main relation between stellar mass and SFR (see fig. 10 in Oteo et al. 2015, and e.g. Rodighiero et al. 2014).

## 2.2 $Ly\alpha$ observations at $z = 2.23$

$Ly\alpha$  observations were conducted at the INT at the Observatorio Roque de los Muchachos on the island of La Palma with a specially designed NB filter for the Wide Field Camera (WFC). This NB filter (NB392,  $\lambda_c = 3918 \text{ \AA}$ ,  $\Delta\lambda = 52 \text{ \AA}$ ) was designed for our survey such that it observes  $Ly\alpha$  emission for all redshifts<sup>1</sup> at which HAEs can be selected with the NB $_K$  filter (see Fig. 1).

The details of the observations, data reduction and calibration are presented in Sobral et al. (in preparation), where we also present the  $Ly\alpha$  luminosity function (LF), and other line-emitters detected in our NB data, such as C IV $_{1549}$  at  $z \approx 1.5$  and [O II] at  $z \approx 0.05$ . For the purpose of this paper, we use the INT observations to measure the  $Ly\alpha$  flux from  $H\alpha$  selected galaxies by creating thumbnail images in NB392. For continuum estimation in COSMOS, we align publicly

available  $U$  and  $B$  bands (from Canada–France–Hawaii Telescope and Subaru, respectively, Capak et al. 2007; McCracken et al. 2010) and measure the flux in these filters at the positions at which the  $H\alpha$  emission is detected. In UDS, we use Canada–France–Hawaii Telescope  $U$ -band data (PI: Almaini & Foucaud) from UKIDSS UDS (Lawrence et al. 2007) and Subaru  $B$ -band data from SXDS (Furusawa et al. 2008).

We converted the  $U$ ,  $B$ , NB $_K$  and  $K$  images to the pixel scale of the INT WFC ( $0.33 \text{ arcsec pixel}^{-1}$ ). The astrometry of COSMOS images is aligned using SCAMP (Bertin 2006), with a reference coordinate system based on *HST* ACS F814W band observations (as in the public COSMOS data; McCracken et al. 2010). The UDS images are aligned to 2MASS (Skrutskie et al. 2006). The accuracy of the astrometry is of the order of  $0.1 \text{ arcsec}$ . We match the full width at half-maximum (FWHM) of the point spread function (PSF) of all images to the FWHM of the NB392 observations (ranging from 1.8 to  $2.0 \text{ arcsec}$ , depending on the particular pointing). The FWHM of reference stars was measured with SExtractor (Bertin & Arnouts 1996), which fits a Gaussian profile to the upper 80 per cent of the light profile from each detected object. For NB392 imaging, we selected reference stars with magnitudes ranging from 16 to 18, resulting in  $\sim 20$  stars per WFC detector. The reference stars in  $U(B)$  are fainter because in  $U(B)$  stars with magnitude  $< 18(19)$  are saturated. For each frame, we find  $\sim 50$  reference stars with magnitudes ranging from 19 to 21. PSF matching was then done by convolving images with a Gaussian kernel. This procedure is based on the PSF matching procedure from the Subaru Suprime-Cam data reduction pipeline (Ouchi et al. 2004).

## 3 MEASUREMENTS

### 3.1 Choice of aperture

Due to resonant scattering of  $Ly\alpha$  photons, the choice of aperture can have an important consequence on the measured  $Ly\alpha$  flux and

<sup>1</sup> Note that we investigate the effect of different filter transmissions between  $Ly\alpha$  and  $H\alpha$  as a function of redshift and the effect of systematic velocity offsets between the lines in Section 3.4.2.



escape fraction, particularly given the evidence of extended Ly $\alpha$  emission for a range of SFGs (e.g. Steidel et al. 2011; Momose et al. 2014 and which we confirm for our sample in Section 7). Previous surveys of LAEs typically used MAG-AUTO photometry with SExtractor to measure Ly $\alpha$  fluxes (e.g. Hayes et al. 2010; Ouchi et al. 2010). However, the measured flux with MAG-AUTO will be dependent on the depth of the NB imaging. As we are measuring Ly $\alpha$  emission for H $\alpha$  selected galaxies at the position of H $\alpha$  detection, it is impossible to perform a similar MAG-AUTO measurement as Ly $\alpha$  selected surveys without uncontrolled bias. This is because we have no a priori knowledge of the optimal aperture to measure Ly $\alpha$ . In fact, we find in Section 5 that most HAEs are undetected in Ly $\alpha$  at the flux limit of our observations. We also note that MAG-AUTO measurements are dependent on the depth, and therefore are not suitable for an optimal comparison as the depth of our survey varies across the field and is different than other surveys.

Due to these considerations, we choose to use fixed diameter aperture measurements for individual sources. An aperture size of 3 arcsec was chosen for the following reasons. First, it corresponds to a radial distance of 12 kpc, which is larger than the exponential scale length of Ly $\alpha$  selected sources at  $z = 2.2$ –6.6 of 5–10 kpc (Momose et al. 2014), and which is also similar to the reference scale used in the study of individual Ly $\alpha$  haloes (Wisotzki et al. 2015; although note that this survey has detected extended Ly $\alpha$  emission up to a radial distance of 25 kpc). Secondly, we find that 3 arcsec aperture magnitudes on the PSF convolved images of the  $U$ ,  $B$ , NB $_K$  and  $K$  band recover similar magnitudes as the 2 arcsec diameter apertures on the original H $\alpha$  images (which typically have a PSF FWHM  $\sim 0.8$  arcsec), with a standard deviation of 0.2 mag. These magnitudes from 2 arcsec aperture measurements are used on most studies of the HAEs from our sample (e.g. Sobral et al. 2013; Oteo et al. 2015). For stacks of subsets of HAEs we vary the aperture, and discuss the difference in Section 7.

### 3.2 Measuring line-fluxes

We use fluxes in NB392,  $U$  and  $B$  band to measure the Ly $\alpha$  line-flux on the positions of the HAEs using dual-mode SExtractor. The NB392 flux is calibrated on  $U$ -band magnitudes of photometrically selected galaxies (see Sobral et al., in preparation), since stars have the strong CaII<sub>3933</sub> absorption feature at the wavelengths of the NB filter. After this calibration, we also make sure that the NB excess ( $U - \text{NB392}$ ) is not a function of the  $U - B$  colour, such that a very blue/red continuum does not bias line-flux measurements. This means that we empirically correct the NB magnitude using:

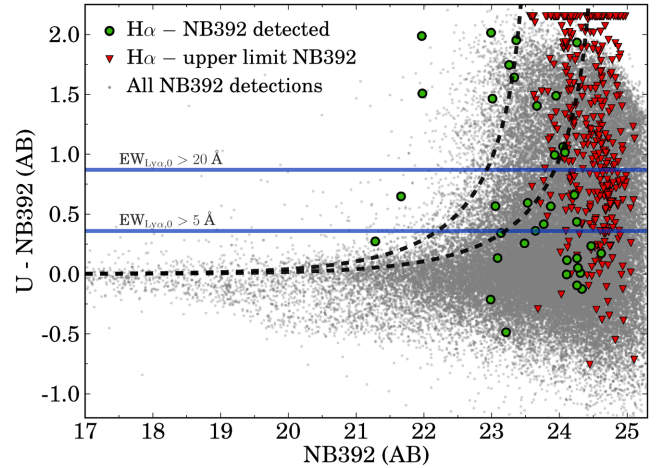
$$\text{NB392}_{\text{corrected}} = \text{NB392} + 0.19 \times (U - B) - 0.09. \quad (1)$$

This correction ensures that a zero NB excess translates into a zero line-flux in NB392. For sources which are undetected in  $U$  or  $B$  we assign the median correction of the sources which are detected in  $U$  and  $B$ , which is +0.02. In the following, we refer to the broad-band  $U$  as BB. Then, with the NB and continuum measurements, the Ly $\alpha$  line-flux is calculated using:

$$f_{\text{Ly}\alpha} = \Delta\lambda_{\text{NB}} \frac{f_{\text{NB}} - f_{\text{BB}}}{1 - \frac{\Delta\lambda_{\text{NB}}}{\Delta\lambda_{\text{BB}}}}. \quad (2)$$

Here,  $f_{\text{NB}}$  and  $f_{\text{BB}}$  are the flux-densities in NB392 and  $U$  and  $\Delta\lambda_{\text{NB}}$  and  $\Delta\lambda_{\text{BB}}$  the filter-widths, which are 52 Å and 758 Å, respectively.

We measure H $\alpha$  line-fluxes as described in Sobral et al. (2013). The relevant NB is NB $_K$  and the continuum is measured in  $K$  band. The excess is corrected with the median correction of +0.03 derived



**Figure 3.** NB excess diagram of the sources in COSMOS and UDS. Grey points show all NB392 detections, where  $U$  has been measured in dual-mode. The green points show the HAEs which are directly detected in the NB392 imaging, with measurements done at the position of the HAEs. The red triangles are upper limits at the positions of the HAEs. The blue horizontal lines show to which rest-frame Ly $\alpha$  EW a certain excess corresponds. Dashed black lines show the excess significance for either the shallowest (left) or deepest (right) NB392 data. Note that some upper limits on the NB392 magnitude are actually weaker than some detections. This is due to variations in the depth of our NB392 observations across the field. Many stars have a negative excess due to the CaII<sub>3933</sub> absorption feature.

from  $H - K$  colours. For an HAE to be selected as a double H $\alpha$ -LAE, we require the  $U - \text{NB392}$  excess to be  $> 0.2$  (corresponding to  $\text{EW}_0 > 4$  Å) and a Ly $\alpha$  excess significance  $\Sigma > 2$  (cf. Bunker et al. 1995; Sobral et al. 2013), see the dashed lines in Fig. 3, which we base on local measurements of the NB and broad-band background in empty 3 arcsec diameter apertures. This relatively low excess significance is only appropriate because we observe pre-selected HAEs. We note, however, that all our directly detected sources are detected with at least  $3\sigma$  significance in the NB392 imaging.

### 3.3 Measuring the Ly $\alpha$ escape fraction

In order to measure the observed fraction of Ly $\alpha$  flux, we need to carefully estimate the intrinsic Ly $\alpha$  line-flux. The intrinsic emission of Ly $\alpha$  due to recombination radiation is related to the H $\alpha$  flux, and scales with the number of ionizing photons per second. Assuming case B recombination, a temperature  $T$  between 5000 and 20 000 K and electron density  $n_e$  ranging from  $10^2$  and  $10^4$  cm $^{-3}$ , the intrinsic ratio of Ly $\alpha$ /H $\alpha$  ranges from 8.1 to 11.6 (e.g. Hummer & Storey 1987). For consistency with other surveys (as discussed by e.g. Hayes 2015; Henry et al. 2015), we assume  $n_e \approx 350$  cm $^{-3}$  and  $T = 10^4$  K, such that the intrinsic ratio between Ly $\alpha$  and H $\alpha$  is 8.7. Therefore, we define the Ly $\alpha$  escape fraction as:

$$f_{\text{esc}} = \frac{f_{\text{Ly}\alpha, \text{obs}}}{8.7 f_{\text{H}\alpha, \text{corrected}}}. \quad (3)$$

In the presence of an AGN, the assumption of case B recombination is likely invalid because e.g. collisional ionization might play a role due to shocks, leading to false estimates of the escape fraction. Among the sample of HAEs, we identify nine X-ray AGN in COSMOS using *Chandra* detections (Elvis et al. 2009), which are all spectroscopically confirmed to be at  $z = 2.23$  (Civano et al. 2012). Eight of these are significantly detected in NB392 imaging.

We exclude these AGN from stacking analyses, but will keep them in our sample for studying individual sources.

Note that since we measure line-fluxes in 3 arcsec apertures,  $f_{\text{esc}}$  is strictly speaking the escape fraction within a radius of 12 kpc. It is possible that the total escape fraction is higher, particularly in the presence of an extended low SB halo due to resonant scattering (see also the discussion from a modeller's point of view by Zheng et al. 2010).

### 3.4 Corrections to measurements

Although our matched NB survey requires less assumptions and uncertain conversions than escape fraction estimates based on UV or other emission-line measurements, we still need to take the following uncertainties/effects into account:

- (i) interlopers in the H $\alpha$  sample (Section 3.4.1)
- (ii) different filter transmissions (Section 3.4.2)
- (iii) dust correction of the observed H $\alpha$  flux (Section 3.4.3)
- (iv) [NII] contributing to the flux in the H $\alpha$  filter (Section 3.4.4).

#### 3.4.1 Interlopers

Our H $\alpha$  sample is selected using photometric redshifts and colour-colour techniques in a sample of emission line galaxies obtained from NB imaging (see Sobral et al. 2013). This means that galaxies with other emission lines than H $\alpha$  can contaminate the sample if the photometric redshift is wrongly assigned (e.g. if the galaxy has anomalous colours). Spectroscopic follow-up shows a 10 per cent interloper fraction, although this follow-up is so far limited to the brightest sources. These interlopers are either dusty low-redshift ( $z < 1$ ) sources, such as Pa $\beta$  at  $z = 0.65$ , or H $\beta$ /[O III] emitters ( $z \sim 3.2$ – $3.3$ ; e.g. Khostovan et al. 2015). For the  $z \sim 3.3$  emitters, the NB392 would only measure noise, as the NB392 filter observes below the Lyman break for higher redshift galaxies and the flux for the low-redshift interlopers is typically much fainter than the NB392 limit. The identified interlopers do not occupy a particular region in the parameter space of the sample of HAEs. There may be small dependences of contamination with galaxy properties, but no trends are seen for our limited follow-up, thus we assume a flat contamination. For stacking, we increase our observed NB392 flux by 10 per cent to account for these interlopers. For individual sources without NB392 detection, we are careful in our analysis as there is the risk of interlopers, even though the fraction is relatively small.

#### 3.4.2 Filter transmissions

While the NB392 and NB $K$  filters are very well matched in terms of redshift coverage, the transmission at fixed redshift varies between H $\alpha$  and Ly $\alpha$ . This means that the measured escape fraction is influenced by the particular redshift of the galaxy and resulting different filter transmissions for H $\alpha$  and Ly $\alpha$ . Furthermore, systematic velocity offsets between Ly $\alpha$  and H $\alpha$  might increase this effect, as it has been found that Ly $\alpha$  is redshifted typically 200 (400) km s $^{-1}$  with respect to H $\alpha$  in Ly $\alpha$  (UV) selected galaxies (e.g. Steidel et al. 2010; McLinden et al. 2011; Kulas et al. 2012; Hashimoto et al. 2013; Erb et al. 2014; Shibuya et al. 2014; Song et al. 2014; Sobral et al. 2015b; Trainor et al. 2015). We test the effect of the different filter transmissions and velocity shifts using a Monte Carlo simulation, similar to e.g. Nakajima et al. (2012). We simulate 1000 000 galaxies with redshifts between the limits of the NB $K$  filter, and with

a redshift probability distribution given by the NB $K$  filter transmission (as our sample is H $\alpha$  selected). Then, we redshift the Ly $\alpha$  line w.r.t. H $\alpha$  with velocity shifts ranging from 0 to 800 km s $^{-1}$ , and fold it through the filter transmission in NB392. Finally, we compute the average relative H $\alpha$ –Ly $\alpha$  transmission. For a zero velocity offset, the average transmission is 20 per cent higher for Ly $\alpha$  than H $\alpha$ , because the NB392 filter is more top-hat like than the NB $K$  filter. Increasing the velocity offset leads to an average lower Ly $\alpha$  transmission, as it is redshifted into lower transmission regions in the right wing of the filter. This effect is, however, very small, as it is constant up to a velocity shift of 400 km s $^{-1}$ , and decreases to 11 per cent for 800 km s $^{-1}$ . Because of this, we decrease the Ly $\alpha$ –H $\alpha$  ratio of stacks and individual sources by 20 per cent. We add the 20 per cent uncertainty of this correction to the error on the escape fractions in quadrature. Spectroscopic follow-up is required to fully investigate the effect of velocity offsets on our measured escape fractions.

#### 3.4.3 Dust attenuation

Even though the H $\alpha$  emission line is at red wavelengths compared to, for example, UV radiation, it is still affected by dust, such that we underestimate the intrinsic H $\alpha$  luminosity. Correcting for dust typically involves a number of uncertainties, such as the shape and normalization of the attenuation curve, the difference between nebular and stellar extinction (e.g. Reddy et al. 2015 and references therein) and the general uncertainties in SED fitting. For consistency with other surveys, we correct for extinction by applying a Calzetti et al. (2000) dust correction, using the estimated extinction,  $E(B - V)_{\text{star}}$ , measurements from the best-fitting SED model from Sobral et al. (2014). Note that we assume  $E(B - V)_{\text{star}} = E(B - V)_{\text{gas}}$ , independent of galaxy property. Recent spectroscopic results at  $z \sim 2$  (e.g. Reddy et al. 2015) indicate that this is reasonable when averaged over the galaxy population, although there are indications that the nebular attenuation is higher than the stellar attenuation for galaxies with high SFR, particularly for galaxies with  $\text{SFR} > 50 M_{\odot} \text{ yr}^{-1}$ . Therefore, if such a trend would be confirmed, our inferred relations between  $f_{\text{esc}}$  may be slightly affected. We discuss this when relevant in Sections 6 and 8.2.1.

When stacking, we use the median dust correction of the sources included in the stacked sample, which is  $A_{\text{H}\alpha} = 1.0$ . This number has also been used, for example, by Sobral et al. (2013) in order to derive the cosmic star formation rate density, which agrees very well with independent measures. Ibar et al. (2013) showed this median attenuation also holds for a similar sample of HAEs at  $z = 1.47$  by using *Herschel* data.

However, we also investigate how our results change when using the dust correction prescription from Garn & Best (2010), which is a calibration between dust extinction and stellar mass based on a large sample of spectroscopically measured Balmer decrements in the local Universe. This relation between Balmer decrement and stellar mass is shown to hold up to at least  $z \sim 1.5$  (e.g. Sobral et al. 2012; Ibar et al. 2013).

For individual sources, the two different dust corrections explored here can vary by up to a factor of 5, as seen in Table 2. This results in large systemic errors which can only be addressed with follow-up spectroscopy to measure Balmer decrements. Throughout the paper, we add the error on the dust correction due to the error in SED fitting in quadrature to the error of the H $\alpha$  flux, but note that the systematic errors in the dust correction are typically of a factor of 2.

### 3.4.4 [N II] contamination

Due to the broadness of the NB<sub>K</sub> filter used to measure H $\alpha$ , the adjacent [N II] emission line doublet contributes to the observed line-flux. We correct for this contribution using the relation from Sobral et al. (2012), who calibrated a relation between [N II]/([N II]+H $\alpha$ ) and EW<sub>0,H $\alpha$ + [N II]</sub> on Sloan Digital Sky Survey galaxies. More recently, Sobral et al. (2015a) found the relation to hold at least up to  $z \sim 1$ . At  $z = 2.23$ , we use this relation to infer a typical fraction of [N II]/([N II]+H $\alpha$ ) =  $0.17 \pm 0.08$ , which is consistent with spectroscopic follow-up at  $z \sim 2$  (Swinbank et al. 2012; Sanders et al. 2015). We have checked that our observed trends between  $f_{\text{esc}}$  and galaxy properties do not qualitatively depend on this correction – if we apply the median correction to all sources, the results are the same within the error bars. We add 10 per cent of the correction to the error in quadrature. For stacks, we measure the EW<sub>0,H $\alpha$ + [N II]</sub> and apply the corresponding correction, which is consistent with the median correction mentioned here, and we also add 10 per cent of the correction to the error in quadrature.

### 3.5 Definitions of galaxy properties

We compare  $f_{\text{esc}}$  with a range of galaxy properties, defined here. SFRs are computed from H $\alpha$  luminosity, assuming a luminosity distance of 17746 Mpc (corresponding to  $z = 2.23$  with our cosmological parameters) and the conversion using a Chabrier (2003) IMF:

$$\text{SFR}(\text{H}\alpha)/(\text{M}_{\odot}\text{yr}^{-1}) = 4.4 \times 10^{-42} L(\text{H}\alpha)/(\text{erg s}^{-1}) \quad (4)$$

where  $L(\text{H}\alpha)$  is the dust-corrected H $\alpha$  luminosity and SFR(H $\alpha$ ) the SFR.

Stellar masses and extinctions ( $E(B - V)$ ) are obtained through SED fitting as described in Sobral et al. (2014). In short, the far-UV to mid-infrared photometry is fitted with Bruzual & Charlot (2003) based SED templates, a Chabrier (2003) IMF, exponentially declining star formation histories, dust attenuation as described by Calzetti et al. (2000) and a metallicity ranging from  $Z = 0.0001$  to 0.05. While we use a mass defined as the median mass of all fitted models within  $1\sigma$  of the best fit, we use the  $E(B - V)$  value of the best-fitted model. The errors on stellar masses and extinctions are computed as the  $1\sigma$  variation in the fitted values from SED models that have a  $\chi^2$  within  $1\sigma$  of the best-fitted model. For stellar mass, these errors range from 0.2 dex for the lowest masses to 0.1 dex for the highest masses. The typical uncertainty on the extinction ranges from 0.12 at  $E(B - V) \approx 0.1$  to 0.05 at  $E(B - V) \approx 0.3$ .

The UV slope  $\beta$  (which is a tracer for dust content, stellar populations and escape of continuum ionizing photons; e.g. Dunlop et al. 2012) is calculated using photometry from the observed  $g^+ - R$  colours. These bands were chosen such that there is minimal contribution from Ly $\alpha$  to the  $g^+$  band (the transmission at the corresponding wavelength is  $<5$  per cent), and such that we measure the slope at a rest-frame UV wavelength of  $\sim 1500 \text{ \AA}$ . We also chose to derive the slope from observed colours instead of using the SED fit, as otherwise there might be biases (e.g. the SED-based extinction correction is related to the UV slope). The error in  $\beta$  due to measurement errors in  $g^+$  and  $R$  ranges from typically 0.5 at  $\beta = -2.3$  to 0.3 at  $\beta > 0$ .

## 4 STACKING METHOD

In order to reach deeper Ly $\alpha$  line-fluxes, we use stacking methods to combine observations of our full sample of observed galaxies,

such that the exposure time is effectively increased by a factor of  $\sim 400$ . This however involves some complications and assumptions. For example, we will use the median stacked value, rather than the mean stacked value, such that our results are not biased towards bright outliers. However, our results will still be biased towards the most numerous kind of sources in our sample. Stacking also assumes that all sources are part of a single population with similar properties – which may not always be the case, as indicated by the results in the previous section.

We divide our sample in subsets of various physical properties in Section 6 and study how these stacks compare with the results from individual galaxies. We discuss the effect of varying apertures in Section 7. The errors of the measured fluxes and resulting escape fractions in stacks are estimated using the jackknife method. The errors due to differences in the PSF of the NB and broad-band are added as a function of aperture radius (see Section 4.1). We add all other sources of systematic error (see Section 3.4) in quadrature.

We obtain stacked measurements by median combining the counts in  $1 \times 1 \text{ arcmin}^2$  thumbnails in  $U$ ,  $B$ , NB392, NB<sub>K</sub> and  $K$  bands of the HAEs covered in our INT observations (see Fig. 2). From the stacked thumbnails (as, for example, shown in Fig. 11), we measure photometry in various apertures at the central position (defined by the position of the NB<sub>K</sub> detection. Note that our typical astrometric errors are of the order of  $\sim 0.1 \text{ arcsec}$ , corresponding to  $\sim 1 \text{ kpc}$ ). With this photometry, we obtain line-fluxes for both Ly $\alpha$  and H $\alpha$ . The Ly $\alpha$  flux is corrected using  $U - B$  colours, and we account for the [N II] contribution to the NB<sub>K</sub> flux using the relation with EW from Sobral et al. (2012) (see Section 3.4.4). We also add the error due to differences in the PSF of  $U$  and NB392 to the error of the Ly $\alpha$  flux (see Section 4.1). We apply the median dust correction of the HAEs, which is roughly similar for using the Calzetti or Garn & Best method:  $A_{\text{H}\alpha} = 1.0$  or  $A_{\text{H}\alpha} = 0.86$ , respectively. For our full sample of 488 HAEs, we observe a median stacked Ly $\alpha$  line-flux of  $3.5 \pm 0.3 \times 10^{-18} \text{ erg s}^{-1} \text{ cm}^{-2}$ , and an escape fraction of  $0.3 \pm 0.06$  per cent in 3 arcsec apertures, corresponding to a radial distance to the centre of  $\sim 20 \text{ kpc}$ . The significance of these detections is discussed in Section 7.

The depth of our NB392 observations is inhomogeneous over the full fields (see Fig. 2). We therefore study the effect of limiting our sample based on the depth of the NB392 observations. We find that the photometric errors on the stacked NB392 image are minimized when we only include sources for which the local  $3\sigma$  depth is at least 24.1 AB magnitude, which corresponds to the inclusion of 265 out of the 488 sources. For the remainder of this section, we only include sources which are among these 265. The median SFRs, stellar masses, and dust attenuations of this sample are similar to the average properties of the full sample (see Table 1). The  $3\sigma$  depth of the NB392 stack of these 265 sources is 27.2 AB magnitude. In the case of a pure line and no continuum contributing to the NB392 flux, this corresponds to a limiting line-flux of  $\sim 5 \times 10^{-18} \text{ erg s}^{-1} \text{ cm}^{-2}$ .

### 4.1 Empirical evaluation of different PSF shapes

The NB and broad-band observations are taken with different telescopes, cameras and at different observing sites and under different conditions. Therefore, even though we match the PSF FWHM of all images, the actual shape of the PSF might vary between NB and broad-band. This might artificially influence the SB profile of line-emission estimated from the difference between the two bands. This becomes particularly important when we study stacked images of over 300 sources, where errors on the per cent level might dominate the measured signal.



**Table 1.** Numbers and median properties with  $1\sigma$  deviations of the sample of H $\alpha$ -LAE with and without AGN. We also show the upper limits on the galaxies that are not detected in Ly $\alpha$ , which is the comparison sample. Note that these are the median upper limits. H $\alpha$  and Ly $\alpha$  fluxes are the values observed in 3 arcsec apertures. For completeness, we show the subsample of SFGs that we used for stacks (these are selected based on the depth of Ly $\alpha$  observations). The masses are derived from SED fitting (Sobral et al. 2014), which also gives H $\alpha$  attenuation based on the stellar extinction and the Calzetti et al. (2000) law (see Section 3.5). The H $\alpha$  attenuation from Garn & Best (2010) is based on a calibration between dust, SFR and mass. The total sample consists of 488 HAEs, with a stacked median escape fraction of  $0.9 \pm 0.1$  per cent (for 3 arcsec diameter apertures), which is lower than the median escape fraction of individually detected source, because for individual sources we are observationally biased towards high escape fractions. \*This escape fraction is likely wrong, as in AGN there is likely a departure from case B recombination due to shocks. We still show this for comparison, indicating that Ly $\alpha$  is typically bright for AGN.

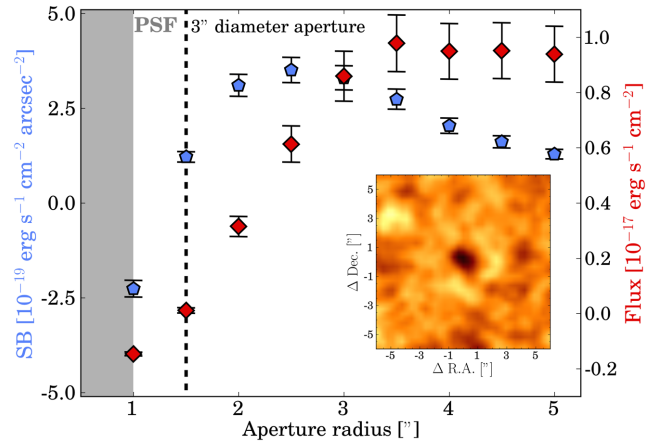
H $\alpha$ sample	Nr.	$\langle f_{H\alpha} \rangle$ [ $10^{-16}$ erg s $^{-1}$ cm $^{-2}$ ]	$\langle f_{Ly\alpha} \rangle$ [ $10^{-16}$ erg s $^{-1}$ cm $^{-2}$ ]	$\log_{10}(M_{\text{star}})$ [ $M_{\odot}$ ]	$A_{H\alpha, \text{Calzetti}}$ [mag]	$A_{H\alpha, \text{Garn\&Best}}$ [mag]	$f_{\text{esc}}$ [ per cent]
SFG with Ly $\alpha$	12	$0.5 \pm 0.3$	$0.7 \pm 0.5$	$10.3 \pm 0.8$	$0.83 \pm 0.4$	$1.11 \pm 0.4$	$10.8 \pm 1.3$
SFG no Ly $\alpha$	468	$0.4 \pm 0.3$	$<0.7$	$9.9 \pm 0.7$	$0.83 \pm 0.5$	$0.86 \pm 0.4$	$<20.1$
AGN with Ly $\alpha$	5	$1.3 \pm 0.4$	$3.6 \pm 1.5$	$10.8 \pm 0.4$	$0.50 \pm 0.3$	$1.55 \pm 0.3$	$12.8 \pm 1.4^*$
SFGs for stacks	265	$0.4 \pm 0.3$	$0.1 \pm 0.01$	$9.9 \pm 0.7$	$1.0 \pm 0.4$	$0.86 \pm 0.4$	$0.9 \pm 0.1$

We empirically evaluate the differences in NB and broad-band PSF by performing the following sanity check: we first select line-emitters in NB $_K$  imaging, which: (i) are *not* selected as HAEs, (ii) are *not* selected as higher redshift line-emitters, or (iii) do *not* have a photometric redshift  $> 1$  (from Ilbert et al. 2009). With this sample, we ensure that the NB392 photometry should measure (relatively flat) continuum by removing a handful of sources with an emission line in NB392. This leaves us with 245 sources, which have a similar NB $_K$  magnitude distribution as the HAE sample. The  $U$ ,  $B$  and NB392 images are stacked in the exact same way as we treat the HAEs. This is used to measure the resulting line-flux and SB profile of the stack in the NB392 band (we estimate the continuum from  $U$  and  $B$ ). Although the NB392 was photometrically calibrated to the  $U$  band in 3 arcsec diameter apertures, we detect a small residual signal with a typical SB profile of central absorption (with SB  $\sim -2 \times 10^{-19}$  erg s $^{-1}$  cm $^{-2}$  arcsec $^{-2}$ ; see Fig. 4), and peaking at a radial distance of 2.5 arcsec (with a SB of  $\sim 4 \times 10^{-19}$  erg s $^{-1}$  cm $^{-2}$  arcsec $^{-2}$ ). We note that at the radial aperture of 1.5 arcsec, which we used for our calibration, the integrated flux signal is consistent with zero (see Fig. 4). Corrections therefore only need to be applied for other aperture radii and SB profiles. For individually detected Ly $\alpha$  sources, the residual signal is at the 1–10 per cent flux level, but for stacks, it can be more important. The origin of this residual signal is likely because of differences in the inner part of the PSF, similarly as those reported by e.g. Momose et al. (2014). The uncertainty in our astrometry is of the order of 0.1 arcsec and therefore likely less important.

We thus conclude that the differences in PSF shapes of broad-band and NB have a small effect on stacked measurements, but we still take it into account by correcting all SB profiles and any aperture measurements at values other than 3 arcsec. We add the residual flux to the error of the total flux in quadrature.

## 5 DIRECT MEASUREMENTS FOR INDIVIDUAL GALAXIES

We directly detect ( $>3\sigma$ ) 43 out of our 488 HAEs in the NB392 imaging, which is a combination of UV continuum and Ly $\alpha$  line (see Table 1). The  $3\sigma$  limit corresponds to limiting Ly $\alpha$  fluxes ranging from  $3.8$  to  $7.4 \times 10^{-17}$  erg s $^{-1}$  cm $^{-2}$  (assuming the typical continuum level of  $0.23 \mu\text{Jy}$ ,  $\sim 25.5$  AB magnitude in the  $U$  band). Out of these robust detections, 17 show a significant Ly $\alpha$  line detection (excess significance  $\Sigma > 2$ ), all in COSMOS. The properties of these sources and their IDs from the HiZELS catalogue (Sobral et al. 2013) are listed in Table 2. The other 26 robust NB392 detec-



**Figure 4.** SB profile (blue) and integrated flux (red) of the stack of the reference sample, which should have a zero line-flux at a 1.5 arcsec aperture radius by construction. We indicate the 1.5 arcsec aperture radius with a dashed black line. The inset figure shows the 2D image obtained by subtracting the BB from the NB. We find a small residual signal which has central absorption and peak at a radial distance of 2.5 arcsec. We attribute this signal due to differences in the PSF of the NB and the broad-band. In the remainder of this paper, we subtract this signal from the SB of individual sources and from stacks and add this subtraction to the error of the total flux in quadrature. The typical signal measured for individually detected HAEs is typically 10–100 times higher than the signal due to the PSF differences, but it is of the same order of magnitude as stacked measurements.

tions are HAEs with strong upper limits on their Ly $\alpha$  flux, as we have detected the UV continuum in the NB392 filter.<sup>2</sup>

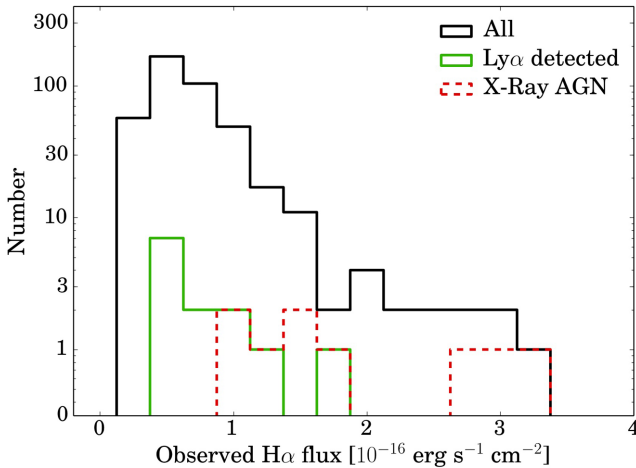
Five of the dual emitters are matched (within 3 arcsec) with an X-ray detection from *Chandra*. From spectroscopy with IMACS and from zCOSMOS (Lilly et al. 2009), these are all classed as BL-AGN. These AGN are among the brightest and most massive HAEs (Fig. 5): all have stellar masses above  $10^{10.5} M_{\odot}$  (Fig. 6), and the fraction of BL-AGN is consistent with the results from Sobral et al. (2016). The ISM conditions surrounding the AGN might lead to other ionizing mechanisms than case B photo-ionization, such

<sup>2</sup> Note that these sources are unlikely higher-redshift interlopers, as the NB392 filter is below the Lyman break at  $z > 3$ , such that a NB392 detection rules out that the source is a  $z \sim 3.3$  H $\beta$ /[O III] or  $z = 4.7$  [O II] emitter (e.g. Khostovan et al. 2015). Moreover, 13 of these have detected [O III], [O II] or both lines.



**Table 2.** Properties of HAEs at  $z = 2.23$  detected at  $3\sigma$  in NB392 imaging and having a  $\text{Ly}\alpha$  emission line. The IDs correspond to the last digits of the full HiZELS IDs, which can be retrieved by placing ‘HiZELS-COSMOS-NBK-DTC-S12B-’ in front of them. The coordinates are measured at the peak of  $\text{NB}_K$  (rest-frame  $R+\text{H}\alpha$ ) emission. The observed line-fluxes, EW, dust extinction and escape fraction are measured with 3 arcsec apertures. The escape fraction ( $f_{\text{esc}}$ ) is computed under the assumptions explained in Section 3.4, thus using  $A_{\text{H}\alpha, \text{C}}$ . <sup>1</sup>: SED: dust correction using SED fitted  $E(B - V)$  and a Calzetti et al. (2000) law. <sup>2</sup>: GB: Garn & Best (2010) dust correction based on stellar mass. <sup>3</sup>: Codes in this column correspond to: 1: X-Ray AGN, 2: [O II] line detected in  $\text{NB}_J$ , 3: [O III] line detected in  $\text{NB}_H$ .

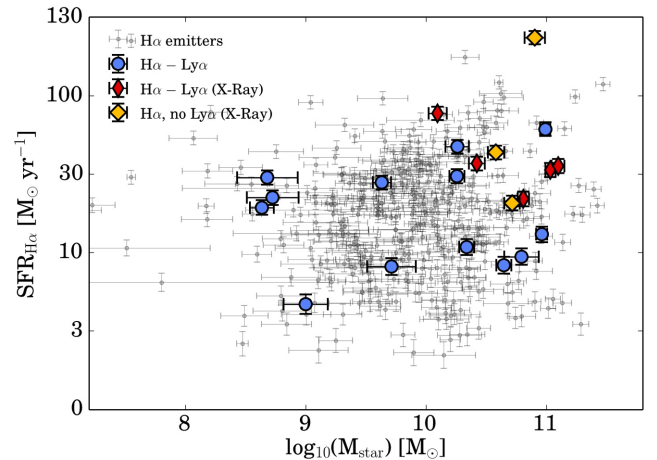
ID	R.A. (J2000)	Dec. (J2000)	$M_{\text{star}}$ $\log_{10}(M_{\odot})$	$f_{\text{Ly}\alpha}$ $\text{erg s}^{-1} \text{cm}^{-2} \times 10^{-16}$	$\text{EW}_{0, \text{Ly}\alpha}$ $\text{\AA}$	$f_{\text{H}\alpha}$ $\text{erg s}^{-1} \text{cm}^{-2} \times 10^{-16}$	$A_{\text{H}\alpha, \text{C}}$ SED <sup>1</sup>	$A_{\text{H}\alpha, \text{GB}}$ GB <sup>2</sup>	$f_{\text{esc}}$ per cent	Note <sup>3</sup>
1057	10:00:39.6	+02:02:41.2	$10.6^{+0.1}_{-0.1}$	0.46	82	0.42	0.17	1.43	$10.8 \pm 1.4$	
1073	10:00:44.2	+02:02:06.9	$11.1^{+0.1}_{-0.1}$	0.88	55	1.24	0.50	1.77	$5.1 \pm 0.6$	1
1139	10:00:55.4	+01:59:55.4	$10.8^{+0.1}_{-0.1}$	4.55	63	1.03	0.17	1.56	$43.7 \pm 5.1$	1, 2, 3
1993	10:02:08.7	+02:21:19.9	$8.6^{+0.1}_{-0.1}$	1.14	68	0.26	1.49	0.29	$12.8 \pm 1.5$	3
2600	10:00:07.6	+02:00:13.2	$8.7^{+0.2}_{-0.2}$	1.32	80	0.53	1.00	0.29	$11.4 \pm 1.3$	3
2741	10:01:57.9	+01:54:36.9	$10.1^{+0.1}_{-0.1}$	3.55	14	2.24	0.67	0.99	$9.9 \pm 1.0$	1
4032	10:00:51.1	+02:41:16.9	$11.0^{+0.1}_{-0.1}$	0.46	28	0.34	0.83	1.67	$7.3 \pm 1.1$	
4427	10:01:19.4	+02:07:32.6	$8.7^{+0.2}_{-0.4}$	0.57	12	0.58	1.16	0.29	$3.8 \pm 0.6$	
4459	10:01:43.3	+02:11:15.7	$10.3^{+0.1}_{-0.1}$	1.41	93	0.41	0.50	1.18	$24.8 \pm 3.1$	
4861	10:00:03.3	+02:11:04.4	$9.0^{+0.2}_{-0.2}$	0.31	16	0.28	0.0	0.34	$12.4 \pm 2.3$	
5583	10:01:59.6	+02:39:32.7	$10.8^{+0.2}_{-0.1}$	1.73	244	0.41	0.50	1.54	$30.4 \pm 3.8$	
5847	10:01:12.2	+02:53:25.9	$10.3^{+0.1}_{-0.1}$	0.70	60	1.03	0.50	1.11	$4.9 \pm 0.5$	
7232	10:01:05.4	+01:46:11.6	$10.3^{+0.1}_{-0.1}$	0.42	15	0.74	1.33	1.11	$1.9 \pm 0.3$	
7693	09:59:49.6	+01:50:24.7	$9.6^{+0.1}_{-0.1}$	0.75	10	0.84	0.67	0.65	$5.5 \pm 0.8$	
7801	10:02:08.6	+01:45:53.6	$10.4^{+0.1}_{-0.1}$	2.38	5	1.35	0.50	1.24	$12.8 \pm 1.4$	1
9274	10:00:26.7	+01:58:23.0	$11.0^{+0.1}_{-0.1}$	5.06	142	1.54	0.13	1.72	$32.4 \pm 3.6$	1, 2
9630	10:02:31.3	+01:58:16.5	$9.7^{+0.2}_{-0.2}$	0.61	29	0.44	0.0	0.70	$16.1 \pm 2.2$	



**Figure 5.** Histogram of  $\text{H}\alpha$  fluxes in our galaxy sample at  $z = 2.23$ . AGN are typically found among the brightest HAEs, and are also typically detected in NB392, such that they are either bright in the UV continuum or  $\text{Ly}\alpha$  or both. We also show the distribution of  $\text{Ly}\alpha$  detected SFGs. It can be seen that not all brightest HAEs have been detected, indicating very low escape fractions or interlopers. On the other hand, some very faint HAEs are still detected in  $\text{Ly}\alpha$ , indicative of high  $\text{Ly}\alpha$  escape fractions.

as shocks, making it more challenging to measure the  $\text{Ly}\alpha$  escape fraction as the intrinsic  $\text{H}\alpha$ - $\text{Ly}\alpha$  changes.

Fig. 5 shows the distribution of  $\text{H}\alpha$  fluxes of our observed sample, of the AGN and also of the star-forming sources directly detected in  $\text{Ly}\alpha$ . Whether a source is detected in  $\text{Ly}\alpha$  does not clearly correlate with  $\text{H}\alpha$  flux, such that even very faint HAEs are detected. These very faint sources generally have high  $f_{\text{esc}}$ , although we note that it



**Figure 6.**  $\text{SFR}(\text{H}\alpha)$  versus stellar mass for the observed HAEs. We obtained the SFR from dust-corrected  $\text{H}\alpha$  and stellar mass from SED fitting (see Section 3.5). We show the position of sources with and without  $\text{Ly}\alpha$ , AGN with  $\text{Ly}\alpha$  and AGN without  $\text{Ly}\alpha$ . There is no obvious difference in the SFRs or stellar masses between sources with or without  $\text{Ly}\alpha$ . Since it is easier to observe  $\text{Ly}\alpha$  for galaxies with higher SFR (at a given  $f_{\text{esc}}$ ), this already indicates that  $f_{\text{esc}}$  is higher for galaxies with low SFR. Note that the SFR for the AGN is likely to be overestimated as AGN activity also contributes to the  $\text{H}\alpha$  flux.

is possible that these sources are detected at a redshift where the transmission in the  $\text{H}\alpha$  filter is low. In the remainder of the paper, we will use the sample of 17 dual  $\text{H}\alpha$ -LAEs for direct measurements of  $f_{\text{esc}}$ , and use upper limits for the other 471 HAEs.

### 5.1 Ly $\alpha$ properties of dual Ly $\alpha$ -HAEs

After excluding the X-ray AGN, we find 12 robust dual Ly $\alpha$ -HAEs, which would translate in 2.5 per cent of our SFGs being detected in Ly $\alpha$  down to our flux limit. However, if we only select the subset of HAEs with the deepest Ly $\alpha$  observations (194 SFGs, of which eight have Ly $\alpha$ ), we find a fraction of 4.1 per cent. This is comparable to Oteo et al. (2015), who found a fraction of 4.5 per cent and lower than the 10.9 per cent of Hayes et al. (2010), whose Ly $\alpha$  observations are a factor of 6 deeper and H $\alpha$  sample consists of fainter sources (by a factor of 7), but covers a volume which is  $\sim 80$  times smaller than our survey.

Comparing our 12 H $\alpha$ -LAEs to HAEs without Ly $\alpha$ , we find that there are no clear differences in the SFR- $M_{\text{star}}$  plane (Fig. 6). Since it is easier to observe Ly $\alpha$  for galaxies with higher SFR (at a given  $f_{\text{esc}}$ ), this already indicates that  $f_{\text{esc}}$  is higher for galaxies with low SFR. We show the stellar masses, observed H $\alpha$  and Ly $\alpha$  fluxes and dust attenuations of the Ly $\alpha$  detected sources in Table 2.

The median, dust-corrected H $\alpha$  luminosity for the 12 SFGs with Ly $\alpha$  is  $4.5 \times 10^{42}$  erg s $^{-1}$  (corresponding to a SFR of  $20 M_{\odot}$  yr $^{-1}$ ), and median stellar mass of  $1.8 \times 10^{10} M_{\odot}$ , such that they have specific SFRs which are typical to the sample of HAEs (see Fig. 6). The median Ly $\alpha$  luminosity is  $2.8 \times 10^{42}$  erg s $^{-1}$  and the Ly $\alpha$  escape fraction for these galaxies ranges from  $1.9 \pm 0.3$  to  $30 \pm 3.8$  per cent (although we note this does not take the uncertainty due to the filter transmission profiles into account, nor a statistical correction.). The  $\text{EW}_0(\text{Ly}\alpha)$  ranges from 10 to 244 Å. If we apply an  $\text{EW}_0(\text{Ly}\alpha)$  cut of  $>25$  Å (similar to the selection of LAEs at high redshift; e.g. Matthee et al. 2014 and references therein), seven out of 194 star-forming HAEs with deepest Ly $\alpha$  observations are recovered as LAEs, with luminosities  $\sim 2\text{--}8 \times 10^{42}$  erg s $^{-1}$ .

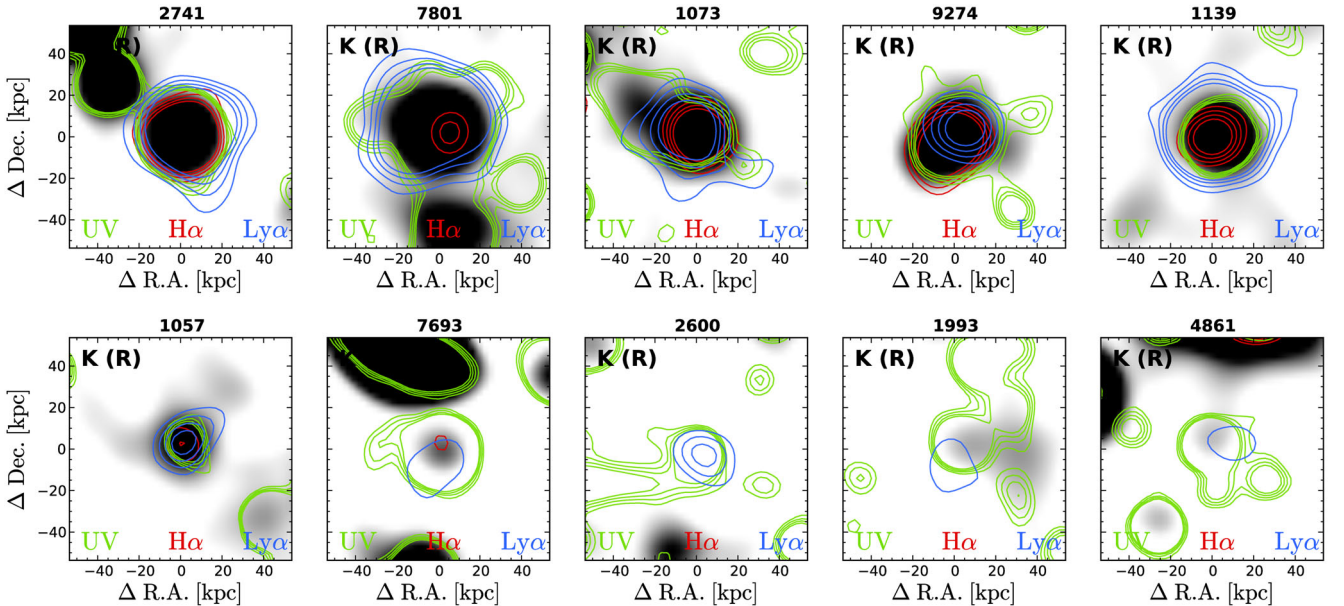
#### 5.1.1 Morphology

In Fig. 7, we compare the Ly $\alpha$  SB with rest-frame UV,  $R$  and H $\alpha$  from *HST* ACS F814W (Koekemoer et al. 2007),  $K$  and NB $\kappa$  (continuum corrected with  $K$ ), respectively. In order to be comparable, the PSF of the *HST* images is matched to that of the NB392 imaging on the INT, using a convolution with a Gaussian kernel. As the PSF of our INT imaging is 1.8–2.0 arcsec, this is a major limitation. However, for the most significantly detected sources (in Ly $\alpha$ ), it is still possible to study differences qualitatively. Even though there is ground-based  $I$ -band data available, we use *HST* data because those are deeper.

The sources with IDs 2741, 78011073, 9274 and 1139 (see Table 2 for more information), shown in the first row, are all AGN with mostly symmetrical Ly $\alpha$  morphology. Compared to the UV, IDs 7801 and 1139 show evidence for extended Ly $\alpha$  emission, while this is more evident when Ly $\alpha$  is compared to H $\alpha$ . Note that the Ly $\alpha$  image is typically the shallowest, and that the outer contours of Ly $\alpha$  therefore typically represent a higher fraction of the peak flux than the UV contours.

The sources in the second row are undetected in the X-ray, and therefore classed as SFGs. 1057 is relatively massive ( $M_{\text{star}} = 10^{10.6} M_{\odot}$ ), and has  $f_{\text{esc}}$  of  $10.8 \pm 1.6$  per cent. 7693 has an intermediate mass of  $\sim 10^{9.5} M_{\odot}$  and escape fraction of  $5.5 \pm 0.8$  per cent. From comparison with the H $\alpha$  image, it can be seen that Ly $\alpha$  preferentially escapes offset to the south from the galaxy centre, which might be indicative of an outflow.

The sources with IDs 2600, 1993 and 4861, in the last three columns of Fig. 7, are the faintest HAEs for which we detect Ly $\alpha$  directly, such that there are no  $3\sigma$  H $\alpha$  contours due to smoothing the image with the PSF of the Ly $\alpha$  image. These HAEs have



**Figure 7.** Ly $\alpha$  morphologies for the 10 HAEs with most significant Ly $\alpha$  detections. All images (including *HST* ACS F814W) were smoothed to the PSF of the NB392 image and are centred on the position of peak H $\alpha$  emission. The top row shows AGN, and the bottom row SFGs. The sources are ordered by H $\alpha$  flux, decreasing from left to right. Each panel shows a  $100 \times 100$  kpc  $K$  band (which traces rest-frame  $R$ , and thus roughly stellar mass) thumbnail with contours of rest-frame Ly $\alpha$  (blue), UV (green, from ACS F814W) and H $\alpha$  (red). The Ly $\alpha$  image was obtained by subtracting the continuum PSF matched  $U$  band from the NB392 image and correcting using the differential PSF image. The contours show 3, 4, 5 and 6  $\sigma$  levels in each respective filter. For Ly $\alpha$ , the  $3\sigma$  contour corresponds to a SB ranging from  $1.8$  to  $17$  (median  $2.0$ )  $\times 10^{-18}$  erg s $^{-1}$  cm $^{-2}$  arcsec $^{-2}$ . The UV data are typically 3 mag deeper than the Ly $\alpha$  data. Among the AGN, IDs 7801 and 1139 show evidence for extended Ly $\alpha$  emission. For the SFGs there is little convincing evidence for extended emission at the reached SB limit. The Ly $\alpha$  emission of source 1993 appears to be offset from the peak UV emission. The faintest HAEs have no  $3\sigma$  contour because of smoothing with the PSF of the Ly $\alpha$  image.

(possibly) little dust, blue UV slopes and  $\text{Ly}\alpha$   $\text{EW}_0 > 30 \text{ \AA}$ , such that they would be selected as LAE. The masses, SFRs and blue UV slopes are consistent with results from typical LAEs (e.g. Nilsson et al. 2009; Ono et al. 2010), and similar to simulated LAEs (e.g. Garel et al. 2012, 2015). The  $\text{Ly}\alpha$  emission for ID 1993 and 4861 appears to be offset from the peak UV emission. This indicates that slit spectroscopy of UV or  $\text{H}\alpha$  selected galaxies might miss significant parts of  $\text{Ly}\alpha$ . Note that we look at the stacked UV,  $\text{Ly}\alpha$ ,  $\text{H}\alpha$  and morphologies of these 12 SFGs and the full sample of SFGs in Section 7.

## 5.2 [O II] and [O III] emission lines

In addition to  $\text{NB}_K$  imaging, the HiZELS survey consists of  $\text{NB}_J$  and  $\text{NB}_H$  imaging in the same fields. These filters are designed such that they also cover the redshifted [O II] (in  $\text{NB}_J$ ) and [O III] (in  $\text{NB}_H$ ) emission at  $z = 2.23$ , similar to e.g. Nakajima et al. (2012) and Sobral et al. (2012). Out of the 488 HAEs, 23 and 70 galaxies are detected in [O II] and [O III], respectively. Two out of the nine AGN are detected as line-emitters in both NBs, two as [O III] emitters, and two AGN as [O II] emitters. The three remaining AGN all have a spectroscopic redshift at  $z = 2.21\text{--}2.23$  (Sobral et al. 2016). This means that all our AGN are either spectroscopically confirmed or have highly accurate photometric redshifts, with emission lines in at least two NBs.

There are two star-forming HAEs with  $\text{Ly}\alpha$  and [O III] (ID 1993 and 2600; see Fig. 7 and Table 2). Compared to all HAEs with detection in [O III] and similar limits on the  $\text{Ly}\alpha$  flux and  $f_{\text{esc}}$ , these galaxies have the lowest mass and SFR ( $\sim 10^{8.6} M_\odot$ ,  $\text{SFR} < 8 M_\odot \text{ yr}^{-1}$ ), most extreme UV slopes (either bluest, ID 2600, or reddest, ID 1993). They have  $f_{\text{esc}} \sim 10$  per cent and [O III]  $\text{EW}_0 \sim 100 \text{ \AA}$ , resembling local Green pea galaxies (e.g. Henry et al. 2015). Their dust corrections are uncertain, because the two methods described in Section 3.4.3 give a factor of 3–5 difference depending on the method. Compared to the other  $\text{Ly}\alpha$  detected HAEs, these galaxies have the highest  $\text{Ly}\alpha$  EWs. ID 2600 has very high  $\text{H}\alpha$   $\text{EW}_0$  of  $\sim 960 \text{ \AA}$ , similar to the most extreme emission line galaxies (e.g. Amorín et al. 2015), while the  $\text{Ly}\alpha$   $\text{EW}_0$  is  $80 \text{ \AA}$ . ID 1993, however, has a relatively low  $\text{H}\alpha$   $\text{EW}_0$  of  $70 \text{ \AA}$ , and  $\text{Ly}\alpha$   $\text{EW}_0$  of  $68 \text{ \AA}$ . We note that if we use the Garn & Best (2010) dust correction, these two galaxies would have a factor of 2–3 higher  $f_{\text{esc}}$ .

## 6 CORRELATIONS BETWEEN $\text{Ly}\alpha$ ESCAPE AND GALAXY PROPERTIES

We use our sample of dual  $\text{Ly}\alpha$ -HAEs and upper limits for the other 468 star-forming HAEs to search for potential correlations between galaxy properties and the  $\text{Ly}\alpha$  escape fraction, shown in Figs 8 and 9. Our sample is compared with the  $\text{H}\alpha$  selected sample of Oteo et al. (2015) and the spectroscopically,  $\text{Ly}\alpha$ -selected sample of Song et al. (2014). The main difference with our sample is that it has deeper  $\text{Ly}\alpha$  observations than Oteo et al. (2015) and spans a wider range of galaxy properties. The major difference in respect to Song et al. (2014) is that their sample is selected on strong  $\text{Ly}\alpha$ , and therefore biased towards sources with high  $f_{\text{esc}}$ .

In addition to studying correlations for individual sources, we also stack different subsets of galaxies. As described in Section 4, we will limit ourselves to the HAEs with the deepest  $\text{NB}_{392}$  observations. We divide our sample by  $\text{SFR}(\text{H}\alpha)$ , stellar mass, dust extinction and UV slope and ensure that our results do not depend on our particular

choice of bin limits and width by perturbing both significantly. The benefit from studying correlations in bins of galaxies is that the results are less dependent on the systemic uncertainties in the dust corrections. While the systematic difference in dust corrections can be up to a factor of 5 for individual sources, the differences are much smaller over a statistical sample (compare for example Tables 1 and 2).

### 6.1 Varying $\text{SFR}(\text{H}\alpha)$ and mass

In the top row of Fig. 8 we show that  $f_{\text{esc}}$  is anti-correlated with SFR. This is seen both in the individual sources and in stacks. As both  $f_{\text{esc}}$  and SFR involve the  $\text{H}\alpha$  measurement, we naively expected to find an anti-correlation between the two. Yet, the quantitative behaviour of this trend is not a priori trivial due to complex  $\text{Ly}\alpha$  radiative transport. The qualitative trend is not strongly sensitive to the dust correction method applied. By fitting a linear relation to our stacked values (in logarithmic space), we obtain

$$\log_{10}(f_{\text{esc}}/(\text{per cent})) = 1.34^{+0.12}_{-0.12} - 1.32^{+0.10}_{-0.10} \log_{10}(\text{SFR}/(M_\odot \text{ yr}^{-1})) \quad (5)$$

for a 12 kpc aperture radius, with  $\chi^2_{\text{red}} = 1.96$ . At a radius of 24 kpc, the normalization is  $2.43^{+0.15}_{-0.15}$  and slope  $-1.65^{+0.08}_{-0.09}$ , with a  $\chi^2_{\text{red}} = 1.27$ . This means that there is an  $\sim 13\sigma$  anti-correlation between the escape fraction and the SFR (as a slope of zero is rejected at  $13\sigma$ ). According to this relation, a galaxy with a SFR of  $4 M_\odot \text{ yr}^{-1}$  has a typical  $\text{Ly}\alpha$  escape fraction of  $\sim 3.5 \pm 1.8$  per cent (in 12 kpc apertures). We note that this relation probably turns over at lower SFR than we probe (as  $f_{\text{esc}}$  would otherwise reach values  $> 100$  per cent). The top-right panel of Fig. 8 also shows that the  $\text{Ly}\alpha$  escape fraction is higher at larger radii at all SFRs, although the errors become a bit larger.

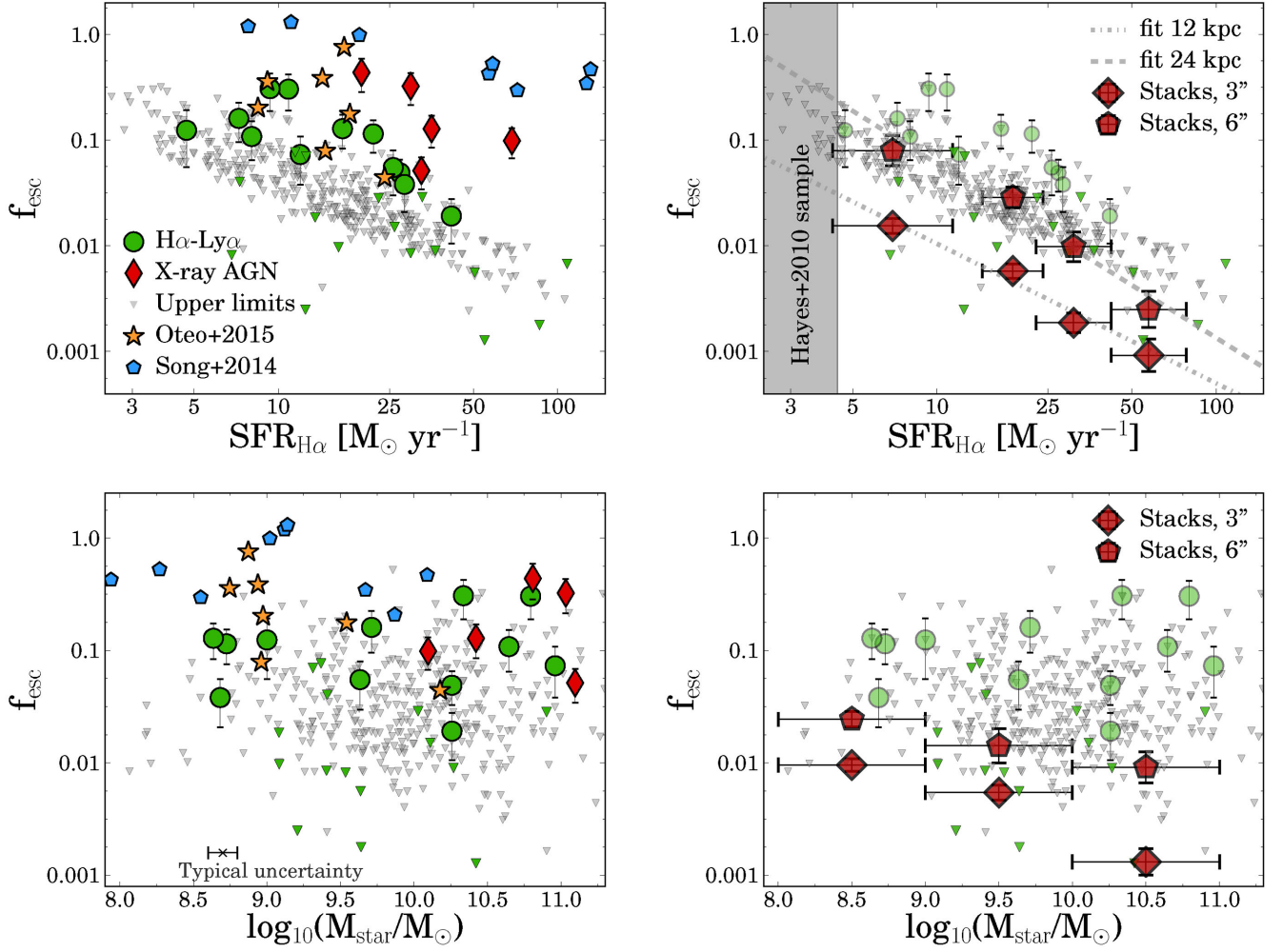
It can be seen that our directly detected sample occupies the region in parameter space which has highest escape fraction at fixed SFR. Compared to the  $\text{Ly}\alpha$  selected sample from Song et al. (2014), we find a stronger anti-correlation between  $f_{\text{esc}}$  and SFR at a lower normalization. This is because  $\text{Ly}\alpha$  selected sources are biased towards high values of  $f_{\text{esc}}$ .

The bottom row of panels in Fig. 8 shows how  $f_{\text{esc}}$  is related to stellar mass. While the stacked values show a weak anti-correlation (although at very low significance), there is no evidence for a trend between  $f_{\text{esc}}$  and stellar mass for the individually detected sources even though our sample clearly extends the probed dynamic range up to higher masses. As massive galaxies would naively be expected to have a lower  $f_{\text{esc}}$ , since they tend to have a higher SFR and are dustier (e.g. Ibar et al. 2013), this means that the  $\text{Ly}\alpha$  escape fraction is not only determined by the scale of a galaxy and it is likely that more subtle processes such as dust and gas dynamics play an important role. Interestingly, the individual detected sources with high stellar mass are at much higher  $f_{\text{esc}}$  than the stacked values, indicating that there is significant scatter in  $f_{\text{esc}}$  in this mass range.

### 6.2 Dust extinction and UV slope

In the top row of Fig. 9 we show how  $f_{\text{esc}}$  is related to dust extinction. For individual sources and for stacks, we find that there is an anti-correlation between  $f_{\text{esc}}$  and dust extinction, although this relation contains significant scatter, caused in part by errors in our measurement of  $f_{\text{esc}}$  and  $E(B - V)$ . With our data, we are able to extend the previous found relation from Hayes et al. (2011) to lower





**Figure 8.**  $\text{Ly}\alpha$  escape fraction versus SFR and stellar mass for galaxies without AGN for individual sources (left panels) and stacks (right panels). The green circles show our directly detected  $\text{H}\alpha$ -LAEs; grey triangles highlight upper limits (green triangles have a UV detection in NB392) and our X-ray identified AGN with  $\text{Ly}\alpha$  are shown in red diamonds. We also add the  $\text{H}\alpha$  selected sample from Oteo et al. (2015) in orange stars and the  $\text{Ly}\alpha$  selected sample from Song et al. (2014) as blue pentagons. Our survey clearly extends the probed parameter space in galaxy properties. Stacked values are shown for two different radial distances to the centre (corresponding to 3 arcsec diameter apertures – diamonds, and 6 arcsec diameter apertures – pentagons). The typical measurement uncertainty in stellar mass is indicated in the bottom left panel. Top row: the left panel shows the SFR obtained from  $\text{H}\alpha$  versus  $f_{\text{esc}}$  of individual sources. Although a correlation is expected by definition, it can be seen that on average, galaxies with a higher SFR have a lower escape fraction. The grey region in the right panel shows the SFRs typical for galaxies in the sample from Hayes et al. (2010), who inferred  $f_{\text{esc}} = 5.3 \pm 3.8$  per cent. Bottom row: escape fraction versus stellar mass. While  $f_{\text{esc}}$  can be relatively high for low-mass galaxies, the stacked results indicate that  $f_{\text{esc}}$  might decrease weakly with increasing stellar mass. The large difference in  $f_{\text{esc}}$  between some massive individual sources and the stacked values indicates that there is likely significant scatter in the values of  $f_{\text{esc}}$  at this mass range.

escape fractions and higher dust extinctions. We fit the following linear relation:

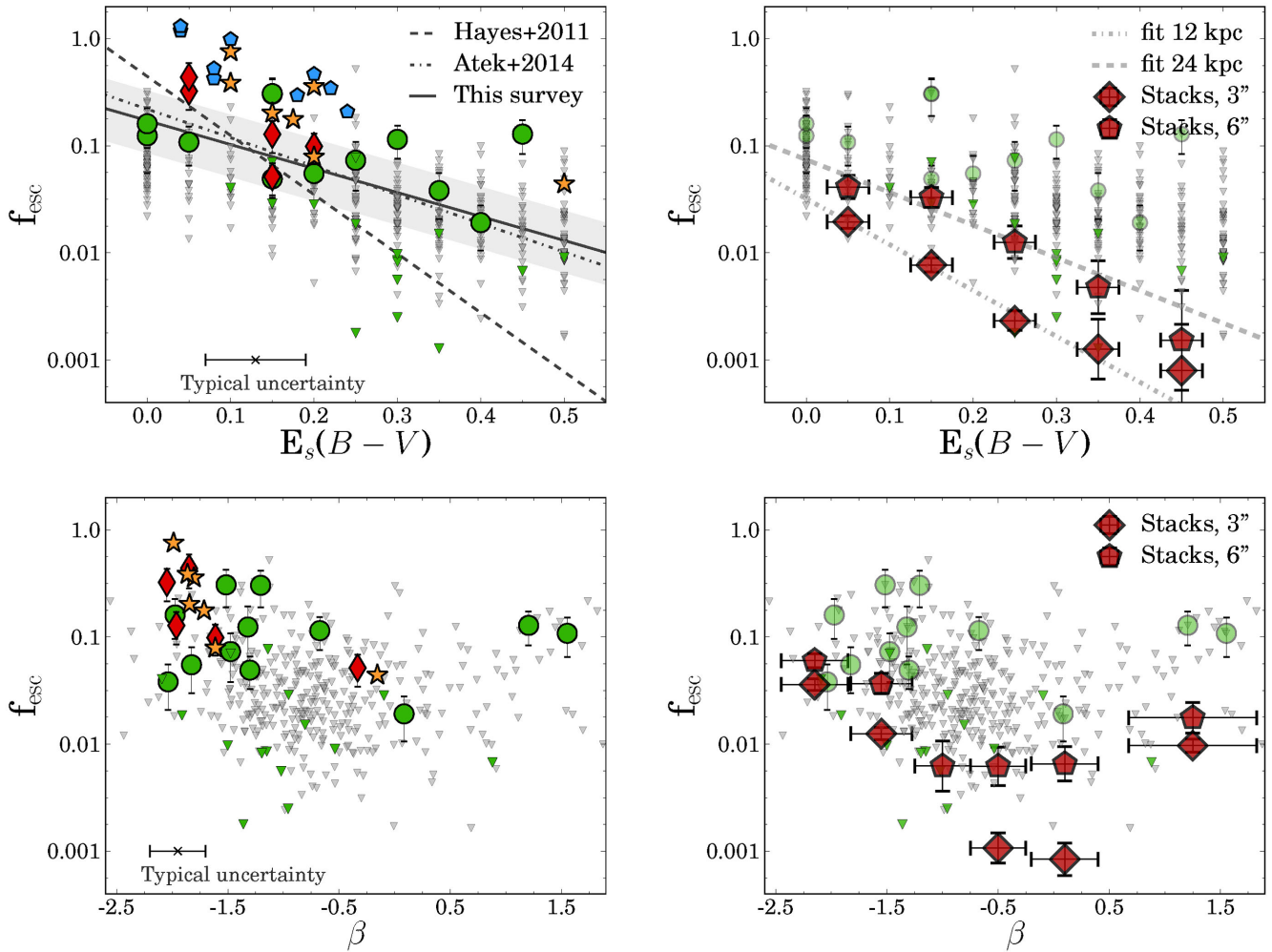
$$f_{\text{esc}} = C \times 10^{-0.4E(B-V)k_{\text{Ly}\alpha}}. \quad (6)$$

This fit is performed by simulating a large grid of normalizations and slopes and computing the  $\chi^2$  for each combination of normalization and slope in  $\log(f_{\text{esc}}) - E(B - V)$  space. Upper limits are taken into account by assigning them an  $f_{\text{esc}}$  of 0.01 per cent and using their upper limit as error. Since a fit to individual galaxies is mostly determined by the directly detected dual  $\text{H}\alpha$ -LAEs, which are biased towards high  $f_{\text{esc}}$ , we also fit to the stacked values.

By minimizing the  $\chi^2$  for individual galaxies, we find  $C = 0.17^{+0.15}_{-0.09}$  and  $k_{\text{Ly}\alpha} = 5.60^{+3.45}_{-2.90}$ , such that a galaxy with  $E(B - V) = 0$  has an escape fraction of 17 per cent. This is lower in normalization (although the errors are significant), and significantly

shallower in slope than the fit from Hayes et al. (2011), who find  $C$  is 0.445 and  $k_{\text{Ly}\alpha} = 13.8$ . The normalization and slope are more similar to the  $z = 0.3$  result from Atek et al. (2014) ( $C = 0.22$ ,  $k_{\text{Ly}\alpha} = 6.67$ ). A possible explanation could be that Hayes et al. (2011) miss dusty galaxies, such that they infer a steeper slope, which would be consistent with the discussion in Oteo et al. (2015). We discuss this further in Section 8.2. For stacks, we find  $C = 0.03^{+0.01}_{-0.01}$  and  $k_{\text{Ly}\alpha} = 10.71^{+0.89}_{-1.01}$  for 12 kpc apertures and  $C = 0.08^{+0.02}_{-0.01}$  and  $k_{\text{Ly}\alpha} = 7.64^{+1.38}_{-1.36}$  for 24 kpc apertures. Our fit to stacked data is less biased towards high values of  $f_{\text{esc}}$  and therefore at a lower normalization. The slope is slightly higher, although still not as high as the slope inferred by Hayes et al. (2011). Similar as seen for the stacks in bins of stellar mass, the individually detected galaxies at highest dust attenuations have much higher  $f_{\text{esc}}$  than the median stack. This means that there is a lot of scatter in the values of  $f_{\text{esc}}$  at the highest dust attenuations.





**Figure 9.** Correlations between  $f_{\text{esc}}$  and dust extinction, rest-frame  $B - V$  colour and  $\beta$ , with symbols as defined in Fig. 8. We indicate the maximum typical uncertainties on extinction and  $\beta$  in the left panels. Top row: the left panel shows that  $f_{\text{esc}}$  is anti-correlated with the dust extinction, although there is significant scatter. This means that at fixed aperture, Ly $\alpha$  preferentially escapes galaxies with low dust content. For comparison, we show the relation at  $z \sim 2-3$  from Hayes et al. (2011) and  $z \sim 0.3$  from Atek et al. (2014). Our best-fitted relation (to detections and upper limits; see text) resembles more than that of local galaxies. The grey band shows the  $1\sigma$  error of the normalization of the fit. Right: stacked values of escape fraction versus dust extinction. A similar trend is seen as for individual sources. However, because stacks are not biased towards high  $f_{\text{esc}}$  values, the normalization is lower. Bottom row: escape fraction versus  $\beta$  for individual sources (left), which confirms that there is a bimodal relation between  $f_{\text{esc}}$  and galaxy UV colour, such that either very blue or very red galaxies emit significant Ly $\alpha$ . As seen in stacks in the right panel, this bimodal trend is most clear at the largest apertures.

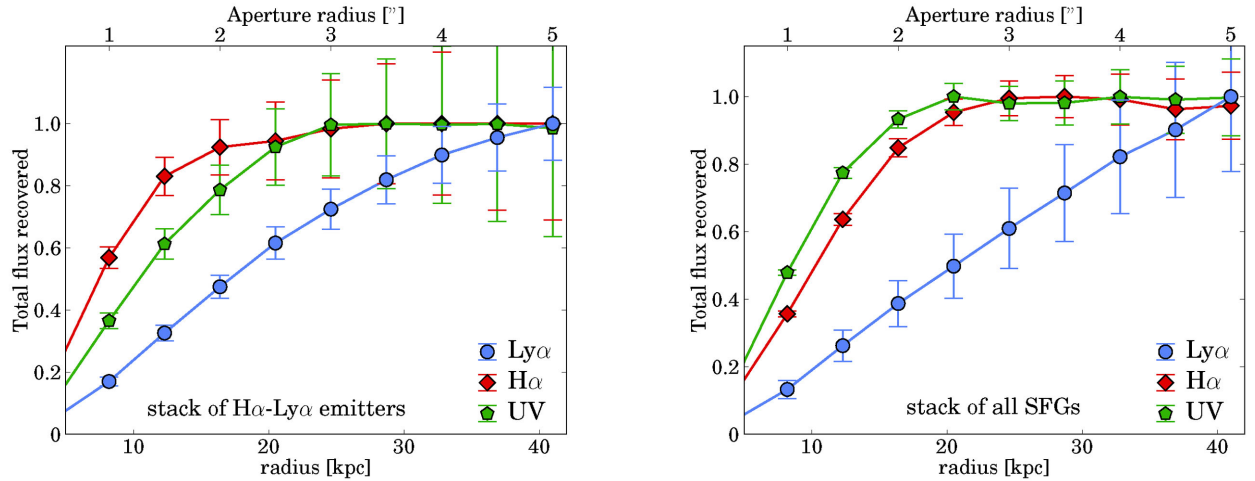
We furthermore note that part of the correlation between  $f_{\text{esc}}$  and  $E(B - V)$  is expected because there is a dust correction in the H $\alpha$  flux, and thus in  $f_{\text{esc}}$ . The fact that the correlation is rather weak (the slope is inconsistent with being zero at only  $1.9\sigma$  and as there is significant scatter) indicates that dust is not the only regulator of Ly $\alpha$  escape, a result already found by Atek et al. (2008) at low redshift. We also note that the trend between  $f_{\text{esc}}$  and  $E(B - V)$  becomes somewhat bimodal when using the Garn & Best (2010) dust correction, meaning that there are galaxies with high dust attenuation and high  $f_{\text{esc}}$ , which is virtually impossible with a Calzetti et al. (2000) dust correction. Galaxies with low  $E(B - V)$  can in this case also have a lower  $f_{\text{esc}}$ , leading to a flattening of the relation.

The bottom row of panels in Fig. 9 shows how  $f_{\text{esc}}$  is related with the UV slope  $\beta$ . As the UV slope is also a tracer of dust attenuation, we also find a tentative anti-correlation with escape fraction for galaxies with  $\beta < 0$ , but for stacks and individual sources, particularly when the HAEs from Oteo et al. (2015) are included.

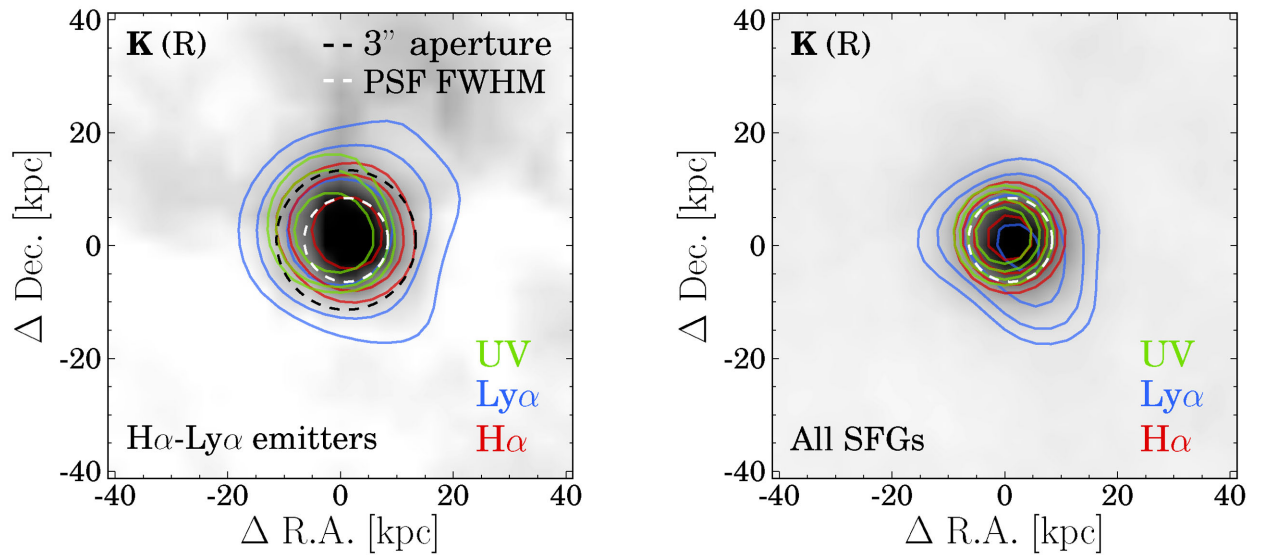
Surprisingly, the trend between  $f_{\text{esc}}$  and  $\beta$  seems to reverse for redder galaxies, leading to a bimodal relation which is particularly seen in measurements of  $f_{\text{esc}}$  with 3 arcsec apertures. The (maximum) typical error of  $\beta$  is indicated in the bottom left panel of Fig. 9 and is sufficiently small to exclude measurement errors of  $\beta$  as the source of the observed bimodal behaviour. The bimodal behaviour is also seen when correcting for dust using the Garn & Best (2010) prescription.

## 7 EXTENDED EMISSION

As Ly $\alpha$  is a resonant emission line, scattering due to neutral hydrogen leads to a diffusion process similar to a random walk, which results in a lower SB. Therefore, it is likely that in the presence of extended H I, Ly $\alpha$  emission will be more extended than the UV. Although our relatively large PSF FWHM and limited depth in NB392 ( $\sim 1.8$  arcsec and  $\sim 24$  AB magnitude, respectively) limit the study of extended Ly $\alpha$  emission at low SB, we can test how our



**Figure 10.** Fraction of the total retrieved flux (defined as the maximum flux within 5 arcsec radius) as a function of aperture for the UV, H $\alpha$  and Ly $\alpha$  and stack of direct detected H $\alpha$ -LAEs (left) and stack of all SFGs (right). It can be seen that while H $\alpha$  and the UV (from PSF convolved *HST* F814W imaging) quickly reaches the total flux in both cases, the Ly $\alpha$  flux continues to increase. For the direct detected sources, the Ly $\alpha$  flux increases somewhat more rapidly because sources were selected on bright central Ly $\alpha$ . These sources also have more compact H $\alpha$  than UV emission. For the stack of all SFGs, Ly $\alpha$  continues to increase (although the errors are significant), such that deeper observations are required to test whether we have captured the full extent of Ly $\alpha$ .



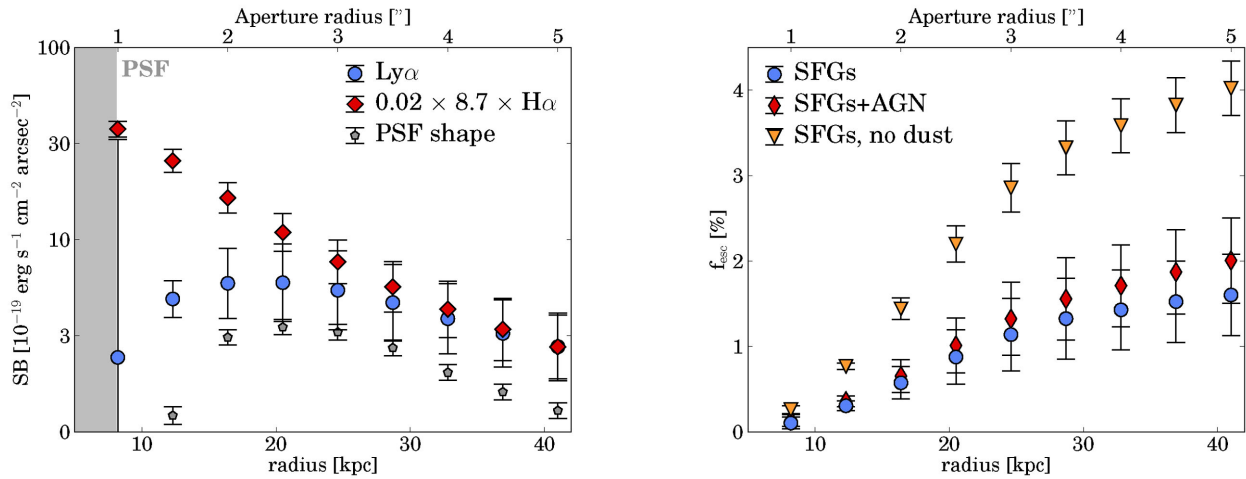
**Figure 11.** Thumbnails of the stacks of the 12 H $\alpha$ -LAEs (left) and full sample of SFGs (right). The background images show the *K* band (where we removed the contribution from H $\alpha$  using NB $_K$ ), which corresponds to rest-frame *R*. The Ly $\alpha$  emission is shown in blue contours, H $\alpha$  emission is shown in red and UV (from observed F814W, corresponding to  $\sim 2000$  Å rest frame) in green. In the left panel we also indicate the 3 arcsec diameter aperture used for measurements of individual sources in black dashed lines. The PSF FWHM is indicated with a white dashed circle in both panels. The contour levels are normalized to the peak flux in each band. The outer Ly $\alpha$  contour represents  $7.5\sigma$  for the left panel, and  $3\sigma$  for the right panel. This corresponds to 0.37 and 0.52 of the peak flux, respectively. The  $1\sigma$  SB limit in the two panels is  $9.0$  and  $2.5 \times 10^{-19}$  erg s $^{-1}$  cm $^{-2}$  arcsec $^{-2}$ . Other contours correspond to a fraction of 0.5 and 0.75 of the peak flux in the left panel, and 0.6, 0.75 and 0.9 in the right panel. In both panels, it can be seen that H $\alpha$  traces the UV light very well. Ly $\alpha$  is more extended than H $\alpha$  and the UV for both the (biased) stack of direct detect H $\alpha$ -Ly $\alpha$  and the full stack of SFGs, indicative of scattering. Ly $\alpha$  extends up to  $\sim 20$  kpc distance from the centre at the corresponding significances. The stack of the 12 H $\alpha$ -LAEs is extended up to 30 kpc at  $3\sigma$  significance.

measured escape fraction depends on the chosen aperture size. We do this by analysing the stacked images for the (biased) sample of 12 dual H $\alpha$ -Ly $\alpha$  emitting SFGs, and for the stack of the 265 SFGs (see Section 5 and Table 1). We measure both H $\alpha$ , Ly $\alpha$  and the rest-frame UV in apertures ranging from 2 to 10 arcsec in diameter.

As seen in Fig. 10 (and illustrated by Fig. 11), Ly $\alpha$  is significantly more extended than the UV (traced by convolved *HST* F814W imaging) and H $\alpha$  for both stacks. The stack of all SFGs (right panel

of Fig. 11) has extended Ly $\alpha$  emission up to  $\sim 20$  kpc distance from the centre, at  $3\sigma$ . At this significance, the stack of the 12 dual H $\alpha$ -LAEs (left panel of Fig. 11) is extended up to 30 kpc, and is clearly more extended than the aperture that we used for the sources individually.

The growth curves for H $\alpha$  and the UV are similar, quickly growing to the maximum at  $\sim 20$  kpc. Ly $\alpha$ , however, continues to increase. The Ly $\alpha$  flux for the stack of the sample of 12 H $\alpha$ -LAEs



**Figure 12.** Left: SB of our full stack of SFGs for H $\alpha$  and Ly $\alpha$  and we also show the SB for the stack of the PSF control sample (Section 4.1), for which the stack of SFGs is corrected. The SB that we observe is attributed to different shapes of the PSF of the NB and broad-band. This SB difference is added to the error of the Ly $\alpha$  SB. We scaled the H $\alpha$  flux such that it has the same SB as Ly $\alpha$  at a radius of 5 arcsec, corresponding to 41 kpc. Note that this scaling corresponds to an escape fraction of 2 per cent, discussed in Section 7.1. Right: escape fraction of the stack of all sources as a function of radial distance from the centre. We compare stacking sources with and without removal of AGN, and we also show the escape fraction if we do not correct H $\alpha$  for dust. The escape fraction at large radii is slightly higher without the removal of AGN, because the AGN are typically bright in Ly $\alpha$  and the haloes around them have a larger scale.

peaks at  $\sim 4$  arcsec radial apertures, as further increase in the fraction of recovered flux with increasing aperture is within the errors. The  $f_{\text{esc}}$  within this radius,  $\sim 30$  kpc, is  $14.2 \pm 1.9$  per cent. At a radius of 1.5 arcsec (similar to the aperture used for individual sources), we measure a stacked  $f_{\text{esc}}$  of  $7.7 \pm 0.9$  per cent. At this aperture, only  $\sim 50$  per cent of the maximum observed Ly $\alpha$  flux is retrieved. The Ly $\alpha$  flux for the stack of all SFGs also continues to increase up to at least 30 kpc. By fitting a linear relation between the fraction of the total recovered flux and radius, we find that  $r_{90}$  (the radius at which 90 per cent of the flux is retrieved) for our stack of directly detected H $\alpha$ -Ly $\alpha$  is  $31.3 \pm 1.5$  kpc. For the stack of all SFGs,  $r_{90} = 36.0 \pm 3.8$  kpc. For H $\alpha$ , we find values of  $r_{90} = 19.3 \pm 1.6$  kpc and  $r_{90} = 21.0 \pm 0.5$  kpc, respectively.

We show the SB profile of the full stack of SFGs in the left panel of Fig. 12, where we scaled H $\alpha$  such that it has a similar SB as Ly $\alpha$  at 5 arcsec radius, corresponding to an escape fraction of 2 per cent (see also the right panel). The H $\alpha$  profile follows an exponential (as the y-axis is logarithmic), decreasing with increasing radius. While the Ly $\alpha$  profile seems to be more complex, we note that the errors due to the different PSF shapes of broad-band and NB are important, particularly in the centre part. Towards higher radii ( $> 30$  kpc), the Ly $\alpha$  signal from the stack is significantly above the errors due to differences in the PSF, and is thus real. We find no evidence that the integrated Ly $\alpha$  flux does not continue to grow up to 40 kpc distance, indicating that it can be extended up to larger radii if deeper SB limits are reached. This also means that we cannot yet directly infer the total  $f_{\text{esc}}$ .

Comparing the H $\alpha$  and Ly $\alpha$  profiles results in an increasing Ly $\alpha$  escape fraction with increasing aperture (see the right panel of Fig. 12). The  $f_{\text{esc}}$  increases from  $0.3 \pm 0.05$  per cent at 12 kpc distance to  $1.6 \pm 0.5$  per cent at  $\sim 30$  kpc. Note that without dust correction  $f_{\text{esc}}$  is roughly a factor of 2 higher. At the radius of 30 kpc, the Ly $\alpha$  SB is  $\sim 6 \times 10^{-19} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ arcsec}^{-2}$  (see also the left panel), such that the extended emission is detected at  $\sim 2.4\sigma$  (at  $2\sigma$  confidence level, extended emission is seen up to  $\sim 40$  kpc). As can be seen in the right panel of Fig. 10 and illustrated in the right panel

of Fig. 11, Ly $\alpha$  is extended up to  $\sim 20$  kpc at  $3\sigma$  confidence level. At  $2\sigma$  significance, it is extended up to  $\sim 30$  kpc. This means that aperture-based measurements (including slit spectroscopy) might miss parts of Ly $\alpha$  emission, and that IFU's or specially designed NB filters are more suited.

## 8 DISCUSSION

### 8.1 A consensus on the value of $f_{\text{esc}}$

To date, a number of papers have been published on measuring the Ly $\alpha$  escape fraction with different selections of galaxies and methods. Typically, Ly $\alpha$  selected galaxies at  $z > 2$  have resulted in high escape fractions of  $\sim 30$  per cent (e.g. Blanc et al. 2011; Nakajima et al. 2012; Wardlow et al. 2014; Kusakabe et al. 2015; Trainor et al. 2015), even though techniques to estimate the intrinsic Ly $\alpha$  production range from UV to dual NB to FIR and spectroscopy. However, a Ly $\alpha$  selected sample of galaxies is not representative for the full (star-forming) galaxy population and estimates of  $f_{\text{esc}}$  are biased towards high values (as e.g. seen in Fig. 8).

With UV or rest-frame optical emission line selected galaxies, the typical  $f_{\text{esc}}$  at  $z = 2-3$  is around  $f_{\text{esc}} \sim 3-5$  per cent (e.g. Hayes et al. 2010; Kornei et al. 2010; Ciardullo et al. 2014), but  $f_{\text{esc}}$  is found to increase with increasing redshift, up to  $\sim 30$  per cent at  $z = 5.7$  (e.g. Hayes et al. 2011). Our measured median value of  $1.6 \pm 0.5$  per cent of HAEs is lower than the value in these papers in the literature. However, we note that our measurement is the first which is independent of assumptions on the shape of the luminosity function and integration limits. Furthermore, as we will show now, the results are fully consistent with literature results when we account for different selections of galaxies and the different parameter spaces probed by different surveys.

Most samples have either been UV selected or selected with emission lines bluer than H $\alpha$  (e.g. [O III]; Ciardullo et al. 2014), such that they might miss a population of dusty galaxies. This might

cause a bias towards high  $f_{\text{esc}}$  values, as we show here that dustier galaxies typically have a lower  $f_{\text{esc}}$  (see Fig. 9, but note that we also observe significant  $f_{\text{esc}}$  for some dusty galaxies). On the contrary, Cassata et al. (2015) report a low fraction of UV selected galaxies with Ly $\alpha$  emission at  $2 < z < 6$ , corresponding to  $f_{\text{esc}} < 1$  per cent, concluding that the bulk of the Ly $\alpha$  luminosity density is coming from fainter galaxies than the typically UV bright galaxies that were targeted. This is consistent with our results, as these galaxies have a typical SFR of  $\sim 50\text{--}100 M_{\odot} \text{ yr}^{-1}$  (Tasca et al. 2015), and we show that the typical escape fraction at these SFRs is very low ( $< 0.5$  per cent).

The major difference between our survey and the matched H $\alpha$ -Ly $\alpha$  survey at  $z = 2.2$  from Hayes et al. (2010) is that our HAEs probe a larger range in SFRs and stellar masses, because our observations probe a much larger volume ( $\sim 80$  times larger, yet with shallower depth). We have shown in the top row of panel of Fig. 8 that the SFR anti-correlates with the escape fraction. If we compute SFRs from the HAEs from Hayes et al. (2010) as described in Section 3.5, we find that their SFRs are between  $0.4$  and  $4.4 M_{\odot} \text{ yr}^{-1}$ , illustrated as the grey region in the top-right panel of Fig. 8. The difference between the inferred escape fraction of  $5.3 \pm 3.8$  per cent by Hayes et al. (2010) and our stack of SFGs can be fully explained by the different parts of parameter space covered, and the anti-correlation between  $f_{\text{esc}}$  and SFR. This also means that the bulk of Ly $\alpha$  luminosity density at  $z = 2.23$  is not emitted by SFGs of  $> 3 M_{\odot} \text{ yr}^{-1}$  (roughly the lower limit of the H $\alpha$  survey), but by much fainter galaxies. We note that Konno et al. (2015) infer a global  $f_{\text{esc}}$  of  $\sim 1.5$  per cent based on more recent measurements of UV and Ly $\alpha$  luminosity functions.

## 8.2 Dependence on galaxy properties

In Section 6 we have shown that  $f_{\text{esc}}$  generally increases with decreasing SFR and dust attenuation (albeit with significant scatter), and that it shows bimodal behaviour with UV slope. The trends between  $f_{\text{esc}}$  and SFR and dust attenuation are also seen in the local Universe (Hayes et al. 2014). In this subsection we discuss how our inferred trends compare to previous results and speculate on the origin of bimodal relations.

### 8.2.1 Relation with dust attenuation

Other surveys have also established a correlation between  $f_{\text{esc}}$  and dust attenuation, both in the local Universe (Atek et al. 2008, 2014) and at high redshift (Verhamme et al. 2008; Hayes et al. 2010; Kornei et al. 2010; Blanc et al. 2011; Cassata et al. 2015). This trend is in line with the significantly lower observed  $f_{\text{esc}}$  in the low-redshift Universe (e.g.  $\sim 1$  per cent at  $z = 0.3\text{--}1$ ; Deharveng et al. 2008; Cowie, Barger & Hu 2010; Wold, Barger & Cowie 2014), as the dust content of galaxies increases with cosmic time (cf. Hayes et al. 2011).

We find a weak trend between  $f_{\text{esc}}$  and dust attenuation, both for stacks and individual galaxies. As shown in the top row of panels of Fig. 9, our trends between  $f_{\text{esc}}$  and  $E(B - V)$  are significantly shallower than the trend inferred by Hayes et al. (2011), and resembles the trend seen in the local Universe (Atek et al. 2014). However, note that comparison is limited by systematic uncertainties and differences between the methods used to estimate  $f_{\text{esc}}$  (e.g. Atek et al. 2014 use Balmer decrements to estimate the dust uncertainty, while our survey is dependent on SED fitting). While Hayes et al. (2010) only included H $\alpha$  and Ly $\alpha$  selected galaxies in their analysis, Hayes et al. (2011) added the UV selected galaxies at  $z \sim 3$  from Kornei

et al. (2010) in order to have a larger sample. However, the UV selected sample requires additional uncertainties on the intrinsic production of Ly $\alpha$  and is biased towards lower dust attenuations (even though their mass and SFR distribution are similar to our H $\alpha$  selected sample).

We speculate that the prime origin of the discrepancy between our fitted trend and the trends from Hayes et al. (2010, 2011) is the combination that we probe more luminous galaxies with higher masses and SFRs (compared to their H $\alpha$  selected sub-sample) and more dusty galaxies (compared to their UV selected sub-sample), and therefore find a lower normalization. As our parameter space also includes more dusty galaxies with significant  $f_{\text{esc}}$ , the relation is flattened. We also note that we find that at a fixed dust attenuation, galaxies with a relatively higher escape fraction have a higher SFR, and are thus also more likely to not be present in the Hayes et al. (2011) sample (as galaxies with higher SFR are rarer), and leading to a flattening of the relation. At the highest dust attenuations, we observe individual galaxies with  $f_{\text{esc}}$  over two orders of magnitude higher than the typical value for galaxies with similar dust attenuation, indicating that there is a lot of dispersion in the values of  $f_{\text{esc}}$ .

It is interesting to note that we find that the  $E(B - V)$  values of HAEs are not dependent on Ly $\alpha$  EW. The relation we would infer between  $f_{\text{esc}}$  and  $E(B - V)$  for HAEs with Ly $\alpha$  EW $_0 > 25 \text{ \AA}$  would be flatter than the relation inferred for our full sample of HAEs. This is because dusty HAEs with high  $f_{\text{esc}}$  tend to have high Ly $\alpha$  EW $_0$ , which may be a sign of outflows. Note however that we do not include Ly $\alpha$  selected sources with faint H $\alpha$  (and thus high  $f_{\text{esc}}$ ) in this analysis.

There still exist large systematic uncertainties on the method to correct for dust attenuation at high redshift. For example, if the nebular extinction is indeed stronger than the stellar extinction at high SFRs (e.g. Reddy et al. 2015), this may influence our observed trends in Figs 8 and 9. We use the Garn & Best (2010) prescription for dust attenuation to evaluate how such a differential dust correction affects our results, as this prescription is based on stellar mass and thus qualitatively similar to a dust attenuation which varies with SFR. On a source-by-source basis, the  $f_{\text{esc}}$  can vary significantly (see e.g. the difference in attenuations in Table 2); however, the results for most stacks are very similar. The largest changes are seen for the stacks with either the highest or lowest SFRs, stellar masses and  $E(B - V)$  values. We find that the relation between  $f_{\text{esc}}$  and  $E(B - V)_{\text{star}}$  flattens somewhat, which is driven by a higher attenuation (and thus lower  $f_{\text{esc}}$ ) for sources in the bin with lowest  $E(B - V)$  values. The relations between  $f_{\text{esc}}$  and SFR and stellar mass would be steeper. Regardless of the method used to correct for dust, the trend we find between  $f_{\text{esc}}$  and  $E(B - V)$  is relatively weak and there is a large dispersion in the values of  $f_{\text{esc}}$ . This indicates that dust attenuation alone is not the most important regulator of  $f_{\text{esc}}$ .

### 8.2.2 Absence of relations and bimodality?

Additional evidence that dust is not the most important (or at least not the only) regulator of  $f_{\text{esc}}$  is that there exists a significant population of red, dusty LAEs (e.g. Blanc et al. 2011; Guaita et al. 2011; Nilsson et al. 2011; Oteo et al. 2015; Taniguchi et al. 2015 and this survey), particularly in the presence of strong outflows (e.g. Atek et al. 2008; Martin et al. 2015). We find tentative evidence that there are bimodal trends between  $f_{\text{esc}}$  and UV slope. This is supported by the observation that the two reddest galaxies in the sample (IDs 1057 and 1993) have a high  $f_{\text{esc}}$ . These sources are surprisingly different: while 1993 has low mass and high dust content, 1057 is



massive with little dust. These two sources are not atypical; in fact, almost all Ly $\alpha$  detected HAEs are either bluer or redder than the average HAE without Ly $\alpha$  detection (which have  $\beta \sim -1$ ). Apart from having high Ly $\alpha$  EW, we find no galaxy property which is related to having a red UV slope and high escape fraction. We note that this bimodal trend can only be seen for samples which include red, relatively massive objects. The trend is therefore still consistent with the results from Hagen et al. (2016), who find no statistical difference in the UV slopes of Ly $\alpha$  or [O III] selected galaxies.

Fig. 8 shows that there is only a weak trend between  $f_{\text{esc}}$  and mass for stacks, such that high mass galaxies have lower  $f_{\text{esc}}$ . However, there are relatively massive galaxies which are detected with high  $f_{\text{esc}}$ . This means that there is a lot of dispersion in the values of  $f_{\text{esc}}$  at high stellar masses. This is also seen in the galaxies with highest dust attenuation. It is interesting to note that the mass scale at which we find higher  $f_{\text{esc}}$  for individual sources coincides with the point where our H $\alpha$  selection starts picking out galaxies below the main relation between stellar mass and SFR (see also Fig. 4). Because of this, and with observations of local galaxies with strong outflows and Ly $\alpha$  emission in mind, we speculate that this enhanced  $f_{\text{esc}}$  might be related due to outflows of (dusty) gas. Outflows in turn redshift Ly $\alpha$  out of resonance wavelengths, which facilitate the escape of Ly $\alpha$  photons, as for example observed in local galaxies (e.g. Bik et al. 2015; Duval et al. 2015; Herenz et al. 2015; Rivera-Thorsen et al. 2015). This idea can be tested with spectroscopic follow-up.

### 8.3 Extended Ly $\alpha$ emission

There is increasing evidence that SFGs are surrounded by a low SB Ly $\alpha$  halo, as indicated from stacks of UV selected galaxies (Steidel et al. 2011) and LAEs (Matsuda et al. 2012; Feldmeier et al. 2013; Momose et al. 2014). Ly $\alpha$  haloes have also been detected around faint individual Ly $\alpha$  selected galaxies, both locally and at high redshift (Hayes et al. 2013; Caminha et al. 2015; Patrício et al. 2016; Wisotzki et al. 2015). Moreover, luminous LAEs, without clear signs of AGN activity, can also be extended up to at least  $\sim 16$  kpc (e.g. Hayashino et al. 2004; Ouchi et al. 2013; Sobral et al. 2015b).

As Ly $\alpha$  emission is observed to be more extended than the UV, it is usually attributed to resonant scattering (e.g. Zheng et al. 2010; Dijkstra & Kramer 2012; Hayes et al. 2013), although simulations from Lake et al. (2015) indicate that collisional ionization due to cooling flows might also contribute. Very extended Ly $\alpha$  emission on scales of  $\sim 100$  kpc has been observed in Ly $\alpha$  blobs (e.g. Matsuda et al. 2004) and Ly $\alpha$  haloes around radio galaxies (e.g. Rottgering et al. 1995; Venemans et al. 2007; Swinbank et al. 2015), believed to be powered by either star formation or AGN activity (e.g. Geach et al. 2009; Overzier et al. 2013; Ao et al. 2015; Umehata et al. 2015).

We find that the average H $\alpha$  selected SFG at  $z = 2.23$  shows Ly $\alpha$  emission which is more extended than their H $\alpha$  or the UV emission (see Figs 10–12). These galaxies are typically dustier than the sources from Steidel et al. (2011) (with our sample of HAEs having a median  $A_{\text{Ly}\alpha} = 3.0$ , compared to  $A_{\text{Ly}\alpha} = 1.12$  from Steidel et al. 2011). The SB of the stack of all sources is therefore fainter (roughly a factor of 2, at a radial distance of 20 kpc) than the stack by Steidel et al. (2011), but we note that the stack of galaxies with  $E(B - V) \sim 0.1$  ( $A_{\text{Ly}\alpha} = 1.2$ ) has similar Ly $\alpha$  SB values as the stack from Steidel et al. (2011). We also note that our stack of SFGs is fainter in the UV by 0.5 mag, but has a dust corrected SFR of  $\sim 42 M_{\odot} \text{ yr}^{-1}$ , compared to the typical SFR of  $\sim 34 M_{\odot} \text{ yr}^{-1}$  of

the sample from Steidel et al. (2011), because of the higher dust attenuation of our sample.

With the current depth, we cannot measure the maximum extent of Ly $\alpha$  for the stack of all SFGs or the stack of directly detected H $\alpha$ -LAEs (see Fig. 10). By comparing the ratio of  $r_{90}$  of Ly $\alpha$  and H $\alpha$  for both stacks, we find that Ly $\alpha$  radii are larger than H $\alpha$  radii with a factor of 1.6–1.7. These values are comparable to those found for local Ly $\alpha$  analogues from Hayes et al. (2013), although that analysis used  $r_{20}$ , which is undesirable for our sample due to the limitations from the PSF. For these local analogues, the average ratio between the Ly $\alpha$  and H $\alpha$  radii is, however, slightly higher than the ratio for our stack of SFGs, which may indicate that we have not yet observed the full extent of Ly $\alpha$ .

For the stack of directly detected H $\alpha$ -LAEs, we observed an escape fraction of  $f_{\text{esc}} = 14.2 \pm 1.9$  per cent at a radius of 30 kpc, twice the  $f_{\text{esc}}$  as the typical aperture used for photometry of Ly $\alpha$  selected galaxies. Assuming the Calzetti et al. (2000) dust law, we can estimate that attenuation of Ly $\alpha$  photons due to dust is a factor of  $9 \pm 3$  ( $A_{\text{Ly}\alpha} = 2.4 \pm 0.3$ ). Hence, this predicts that  $f_{\text{esc}}$  before dust attenuation is  $100/(9 \pm 3)$  per cent. This indicates that dust attenuation is sufficient to explain that  $f_{\text{esc}}$  is  $\sim 15$  per cent. This may well be the case for the majority of Ly $\alpha$  selected galaxies. However, we note that the sample of directly detected H $\alpha$ -LAEs is obviously biased towards centrally peaked, high central  $f_{\text{esc}}$  Ly $\alpha$  emission, as otherwise they would not have been directly detected in Ly $\alpha$ . It is therefore likely that the typical SFG has a larger Ly $\alpha$  halo with a less peaked central SB. The total inferred  $f_{\text{esc}}$  is roughly a factor of 2 higher than  $f_{\text{esc}}$  measured with 3 arcsec diameter apertures. This means that, if we naively assume that this factor of 2 is constant for all 12 SFGs used in the stack, the total escape fractions of our directly detected SFGs (see Table 2) range from 4 to 60 per cent.

## 9 CONCLUSIONS

We have undertaken a panoramic matched H $\alpha$ -Ly $\alpha$  NB survey to study the Ly $\alpha$  emission from a sample of 488 H $\alpha$  selected SFGs at  $z = 2.23$  and its dependence on radial scale and galaxy properties. Our conclusions are as follows.

(i) Out of the 488 observed HAEs, we detect 43 sources in the NB392 imaging. 17 of these have strong Ly $\alpha$  emission (of which five are X-ray AGN), and we measure the UV continuum for the other 26. We put limits on the escape fractions for these 26 and the other 445 undetected HAEs.

(ii) The observed Ly $\alpha$  escape fraction for individual detected sources ranges from 2 to 30 per cent, and we find that these HAEs probe a wide range of star-forming systems: with masses from  $3 \times 10^8 M_{\odot}$  to  $10^{11} M_{\odot}$  and dust attenuations  $E(B - V) = 0 - 0.5$ . We particularly note that some massive, dusty galaxies are clearly visible in Ly $\alpha$ , while they have no evidence for AGN activity from X-ray observations.

(iii) With matched NB observations in the *J* and *H* band from Sobral et al. (2013), we are also able to detect [O II] and [O III] emission lines for 23 and 70 galaxies, respectively. Two faint Ly $\alpha$  emitting galaxies are detected in [O III]. Remarkably, these have relatively high escape fraction and have the lowest stellar mass and highest Ly $\alpha$  EW from our sample.

(iv) While Ly $\alpha$  morphologies of individual X-ray selected AGN are typically circularly symmetric, tracing the rest-frame UV light, the Ly $\alpha$  morphologies of SFGs are more irregular. Ly $\alpha$  can appear

to be offset from the rest-frame UV and H $\alpha$  emission, and also more extended, although we detect this only significantly for AGN.

(v) Both for the individually detected sources and by stacking, we confirm existing trends between the Ly $\alpha$  escape fraction and SFR and dust attenuation and explore how they are extended to a larger range of parameter space – particularly more massive and dusty galaxies. Ly $\alpha$  escape increases for galaxies with lower SFR and low dust attenuation, albeit with significant scatter. The trend between  $f_{\text{esc}}$  and dust attenuation resembles that observed in local galaxies. The escape fraction shows bimodal behaviour with UV slope:  $f_{\text{esc}}$  increases for the bluest and reddest galaxies and is at a minimum UV slope  $\beta \sim -0.5$ . At the highest masses and dust attenuations, we detect individual galaxies with  $f_{\text{esc}}$  much higher than the typical values from stacking, indicating significant scatter in the values of  $f_{\text{esc}}$  at the highest masses and dust attenuations.

(vi) We interpret apparently contradictory reported numbers of the escape fraction by various studies using our observed trends with galaxy properties. Deeper surveys over a smaller area infer a higher escape fraction as the SFRs of their galaxies are lower in general. As these studies miss rarer galaxies with higher SFR and a higher dust attenuation, the inferred correlation between  $f_{\text{esc}}$  and dust content steepens. Studies based on UV selected galaxies report higher numbers of  $f_{\text{esc}}$  because the galaxies in their sample are biased towards galaxies with little dust, which tend to have higher  $f_{\text{esc}}$ , except if their typical SFR is higher.

(vii) The stack of our directly detected H $\alpha$ -LAEs reveals that Ly $\alpha$  is more extended than the UV and H $\alpha$ , with an extent of at least  $\sim 30$  kpc. For Ly $\alpha$ , the radius at which 90 per cent of the flux is retrieved is  $r_{90} = 31.3 \pm 1.5$  kpc, roughly 1.6 times larger than H $\alpha$ . The median escape fraction for this biased sample increases up to  $14.2 \pm 1.9$  per cent at least. We suggest that the missing Ly $\alpha$  photons can be fully explained by dust absorption. At the typical apertures used for slit spectroscopy (e.g. 1 arcsec), only half of the Ly $\alpha$  flux is recovered, even for the sample which is biased towards centrally peaked Ly $\alpha$  emission. This means that slit spectroscopy might miss a significant fraction of Ly $\alpha$ .

(viii) By stacking our full sample of  $z = 2.23$  SFGs, we find that Ly $\alpha$  extends up to at least 3 arcsec (with  $r_{90} = 36.0 \pm 3.8$  kpc) in radial distance at  $3\sigma$ , corresponding to  $\sim 20$  kpc ( $30$  kpc,  $2\sigma$ ), roughly twice the scale of star formation as traced by H $\alpha$  and by the UV, with a Ly $\alpha$  SB of  $\sim 6 \times 10^{-19}$  erg s $^{-1}$  cm $^{-2}$  arcsec $^{-2}$ . As a result, the Ly $\alpha$  escape fraction of our full sample of HAEs continues to increase with radius, up to at least  $1.6 \pm 0.5$  per cent at 30 kpc distance from the centre. With the current depth, it is not yet possible to constrain the maximum extent of Ly $\alpha$ .

Two important questions regarding the escape of Ly $\alpha$  remain. First, what is the total  $f_{\text{esc}}$  for the typical SFG at  $z = 2.23$ ? Deeper Ly $\alpha$  observations are required to fully answer this by measuring the full extent of Ly $\alpha$ . Secondly, properties which are not studied in this survey are likely equally or more important (e.g. the H I covering fraction and geometry, e.g. Scarlata et al. 2009; H I column density, Henry et al. 2015; outflows, Atek et al. 2008; ionization state, Hayes et al. 2014). From the significant scatter in the relations and from the bimodal behaviour in, for example, the relation between  $f_{\text{esc}}$  and dust content and UV slope, it is clear that multiple mechanisms are likely together responsible for determining the Ly $\alpha$  escape fraction and its relation to extended Ly $\alpha$  emission.

It is crucial to improve the measurements of dust attenuation and to obtain precise measurements of redshifts and velocity offsets between Ly $\alpha$  and H $\alpha$ , as these are responsible for the largest systematic errors which can go up to an order of magnitude when

combined in unfortunate ways. However, the major downside is that this requires large, spectroscopic samples over a wide wavelength regime. Additional deeper Ly $\alpha$  NB observations are required in order to eliminate the biases caused by binning and stacking, and to be able to study the relation between galaxy properties and spatial extent of Ly $\alpha$  on a source-by-source basis.

## ACKNOWLEDGEMENTS

We thank the anonymous referee for constructive comments and suggestions which have improved the quality of this work. JM acknowledges the support of a Huygens PhD fellowship from Leiden University. DS and JM acknowledge financial support from the Netherlands Organization for Scientific research (NWO) through a Veni fellowship, and DS from FCT through a FCT Investigator Starting Grant and Start-up Grant (IF/01154/2012/CP0189/CT0010) and from FCT grant PEst-OE/FIS/UI2751/2014. IO acknowledges support from the European Research Council (ERC) in the form of Advanced Investigator Programme, COSMICISM, 321302. HR acknowledges support from the ERC Advanced Investigator programme NewClusters 321271. IRS acknowledges support from STFC (ST/L00075X/1), the ERC Advanced Investigator programme DUSTYGAL 321334 and a Royal Society/Wolfson Merit Award. APA acknowledges support from the Fundação para a Ciência e para a Tecnologia (FCT) through the Fellowship SFRH/BD/52706/2014.

Based on observations made with the Isaac Newton Telescope (proposals 2013AN002, 2013BN008, 2014AC88, 2014AN002, 2014BN006, 2014BC118) operated on the island of La Palma by the Isaac Newton Group in the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofísica de Canarias. We acknowledge the tremendous work that has been done by both COSMOS and UKIDSS UDS/SXDF teams in assembling such large, state-of-the-art multi-wavelength data sets over such wide areas, as those have been crucial for the results presented in this paper. The sample of HAEs is publicly available from Sobral et al. (2013).

We have benefited greatly from the publically available programming language PYTHON, including the NUMPY, MATPLOTLIB, PYFITS, SCIPY (Jones et al. 2001; Hunter 2007; Van Der Walt, Colbert & Varoquaux 2011) and ASTROPY (Astropy Collaboration et al. 2013) packages, the imaging tools SExtractor, SWARP and SCAMP (Bertin & Arnouts 1996; Bertin 2006, 2010) and the TOPCAT analysis program (Taylor 2005).

## REFERENCES

- Amorín R. et al., 2015, A&A, 578, A105
- Ao Y. et al., 2015, A&A, 581, A132
- Astropy Collaboration et al., 2013, A&A, 558, A33
- Atek H., Kunth D., Hayes M., Östlin G., Mas-Hesse J. M., 2008, A&A, 488, 491
- Atek H., Kunth D., Schaerer D., Mas-Hesse J. M., Hayes M., Östlin G., Kneib J.-P., 2014, A&A, 561, A89
- Bertin E., 2006, in Gabriel C., Arviset C., Ponz D., Enrique S., eds, Astronomical Society of the Pacific Conference Series Vol. 351, Astronomical Data Analysis Software and Systems XV. Astron. Soc. Pac., San Francisco, p. 112
- Bertin E., 2010, SWarp: Resampling and Co-adding FITS Images Together, Astrophysics Source Code Library (ascl:1010.068)
- Bertin E., Arnouts S., 1996, A&AS, 117, 393
- Best P. et al., 2013, Astrophys. Space Sci. Proc., 37, 235
- Bik A., Östlin G., Hayes M., Adamo A., Melinder J., Amram P., 2015, A&A, 576, L13

- Blanc G. A. et al., 2011, *ApJ*, 736, 31
- Boquien M., Buat V., Perret V., 2014, *A&A*, 571, A72
- Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000
- Bunker A. J., Warren S. J., Hewett P. C., Clements D. L., 1995, *MNRAS*, 273, 513
- Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, *ApJ*, 533, 682
- Caminha G. B. et al., 2015, preprint ([arXiv:1512.05655](https://arxiv.org/abs/1512.05655))
- Cantalupo S., Arrigoni-Battaia F., Prochaska J. X., Hennawi J. F., Madau P., 2014, *Nature*, 506, 63
- Capak P. et al., 2007, *ApJS*, 172, 99
- Cassata P. et al., 2015, *A&A*, 573, A24
- Chabrier G., 2003, *PASP*, 115, 763
- Ciardullo R. et al., 2014, *ApJ*, 796, 64
- Civano F. et al., 2012, *ApJS*, 201, 30
- Cowie L. L., Barger A. J., Hu E. M., 2010, *ApJ*, 711, 928
- Deharveng J.-M. et al., 2008, *ApJ*, 680, 1072
- Dijkstra M., 2014, *PASA*, 31, 40
- Dijkstra M., Kramer R., 2012, *MNRAS*, 424, 1672
- Dunlop J. S., McLure R. J., Robertson B. E., Ellis R. S., Stark D. P., Cirasuolo M., de Ravel L., 2012, *MNRAS*, 420, 901
- Duval F. et al., 2015, preprint ([arXiv:1512.00860](https://arxiv.org/abs/1512.00860))
- Elvis M. et al., 2009, *ApJS*, 184, 158
- Erb D. K. et al., 2014, *ApJ*, 795, 33
- Feldmeier J. J. et al., 2013, *ApJ*, 776, 75
- Furusawa H. et al., 2008, *ApJS*, 176, 1
- Garel T., Blaizot J., Guiderdoni B., Schaerer D., Verhamme A., Hayes M., 2012, *MNRAS*, 422, 310
- Garel T., Blaizot J., Guiderdoni B., Michel-Dansac L., Hayes M., Verhamme A., 2015, *MNRAS*, 450, 1279
- Garn T., Best P. N., 2010, *MNRAS*, 409, 421
- Geach J. E., Smail I., Best P. N., Kurk J., Casali M., Ivison R. J., Coppin K., 2008, *MNRAS*, 388, 1473
- Geach J. E. et al., 2009, *ApJ*, 700, 1
- Gronwall C. et al., 2007, *ApJ*, 667, 79
- Guaita L. et al., 2011, *ApJ*, 733, 114
- Hagen A. et al., 2016, *ApJ*, 817, 79
- Hashimoto T., Ouchi M., Shimasaku K., Ono Y., Nakajima K., Rauch M., Lee J., Okamura S., 2013, *ApJ*, 765, 70
- Hayashino T. et al., 2004, *AJ*, 128, 2073
- Hayes M., 2015, *PASA*, 32, 27
- Hayes M. et al., 2010, *Nature*, 464, 562
- Hayes M., Schaerer D., Östlin G., Mas-Hesse J. M., Atek H., Kunth D., 2011, *ApJ*, 730, 8
- Hayes M. et al., 2013, *ApJ*, 765, L27
- Hayes M. et al., 2014, *ApJ*, 782, 6
- Henry A., Scarlata C., Martin C. L., Erb D., 2015, *ApJ*, 809, 19
- Herenz E. C. et al., 2015, preprint ([arXiv:1511.05406](https://arxiv.org/abs/1511.05406))
- Hummer D. G., Storey P. J., 1987, *MNRAS*, 224, 801
- Hunter J. D., 2007, *Comput. Sci. Eng.*, 9, 90
- Ibar E. et al., 2013, *MNRAS*, 434, 3218
- Ilbert O. et al., 2009, *ApJ*, 690, 1236
- Jones E. et al., 2001, *SciPy*: Open source scientific tools for Python, <http://www.scipy.org/>
- Kashikawa N. et al., 2012, *ApJ*, 761, 85
- Khostovan A. A., Sobral D., Mobasher B., Best P. N., Smail I., Stott J. P., Hemmati S., Nayyeri H., 2015, *MNRAS*, 452, 3948
- Koekemoer A. M. et al., 2007, *ApJS*, 172, 196
- Konno A., Ouchi M., Nakajima K., Duval F., Kusakabe H., Ono Y., Shimasaku K., 2015, preprint ([arXiv:1512.01854](https://arxiv.org/abs/1512.01854))
- Kornei K. A., Shapley A. E., Erb D. K., Steidel C. C., Reddy N. A., Pettini M., Bogosavljević M., 2010, *ApJ*, 711, 693
- Kulas K. R., Shapley A. E., Kollmeier J. A., Zheng Z., Steidel C. C., Hainline K. N., 2012, *ApJ*, 745, 33
- Kunth D., Mas-Hesse J. M., Terlevich E., Terlevich R., Lequeux J., Fall S. M., 1998, *A&A*, 334, 11
- Kusakabe H., Shimasaku K., Nakajima K., Ouchi M., 2015, *ApJ*, 800, L29
- Lake E., Zheng Z., Cen R., Sadoun R., Momose R., Ouchi M., 2015, *ApJ*, 806, 46
- Lawrence A. et al., 2007, *MNRAS*, 379, 1599
- Lilly S. J. et al., 2009, *ApJS*, 184, 218
- McCracken H. J. et al., 2010, *ApJ*, 708, 202
- McLinden E. M. et al., 2011, *ApJ*, 730, 136
- Martin C. L., Dijkstra M., Henry A., Soto K. T., Danforth C. W., Wong J., 2015, *ApJ*, 803, 6
- Matsuda Y. et al., 2004, *AJ*, 128, 569
- Matsuda Y. et al., 2012, *MNRAS*, 425, 878
- Matthee J. J. A. et al., 2014, *MNRAS*, 440, 2375
- Matthee J., Sobral D., Santos S., Röttgering H., Darvish B., Mobasher B., 2015, *MNRAS*, 451, 400
- Momose R. et al., 2014, *MNRAS*, 442, 110
- Nakajima K. et al., 2012, *ApJ*, 745, 12
- Nilsson K. K., Tapken C., Möller P., Freudling W., Fynbo J. P. U., Meisenheimer K., Laursen P., Östlin G., 2009, *A&A*, 498, 13
- Nilsson K. K., Östlin G., Möller P., Möller-Nilsson O., Tapken C., Freudling W., Fynbo J. P. U., 2011, *A&A*, 529, A9
- Oesch P. A. et al., 2015, *ApJ*, 804, L30
- Ono Y., Ouchi M., Shimasaku K., Dunlop J., Farrah D., McLure R., Okamura S., 2010, *ApJ*, 724, 1524
- Oteo I., Sobral D., Ivison R. J., Smail I., Best P. N., Cepa J., Pérez-García A. M., 2015, *MNRAS*, 452, 2018
- Ouchi M. et al., 2004, *ApJ*, 611, 660
- Ouchi M. et al., 2008, *ApJS*, 176, 301
- Ouchi M. et al., 2010, *ApJ*, 723, 869
- Ouchi M. et al., 2013, *ApJ*, 778, 102
- Overzier R. A., Nesvadba N. P. H., Dijkstra M., Hatch N. A., Lehnert M. D., Villar-Martín M., Wilman R. J., Zirm A. W., 2013, *ApJ*, 771, 89
- Patrício V. et al., 2016, *MNRAS*, 456, 4191
- Reddy N. A. et al., 2015, *ApJ*, 806, 259
- Rivera-Thorsen T. E. et al., 2015, *ApJ*, 805, 14
- Rodighiero G. et al., 2014, *MNRAS*, 443, 19
- Rosdahl J., Blaizot J., 2012, *MNRAS*, 423, 344
- Röttgering H. J. A., Hunstead R. W., Miley G. K., van Ojik R., Wieringa M. H., 1995, *MNRAS*, 277, 389
- Sanders R. L. et al., 2015, *ApJ*, 799, 138
- Scarlata C. et al., 2009, *ApJ*, 704, L98
- Shibuya T. et al., 2014, *ApJ*, 788, 74
- Skrutskie M. F. et al., 2006, *AJ*, 131, 1163
- Sobral D. et al., 2009, *MNRAS*, 398, L68
- Sobral D., Best P. N., Matsuda Y., Smail I., Geach J. E., Cirasuolo M., 2012, *MNRAS*, 420, 1926
- Sobral D., Smail I., Best P. N., Geach J. E., Matsuda Y., Stott J. P., Cirasuolo M., Kurk J., 2013, *MNRAS*, 428, 1128
- Sobral D., Best P. N., Smail I., Mobasher B., Stott J., Nisbet D., 2014, *MNRAS*, 437, 3516
- Sobral D. et al., 2015a, *MNRAS*, 451, 2303
- Sobral D., Matthee J., Darvish B., Schaerer D., Mobasher B., Röttgering H. J. A., Santos S., Hemmati S., 2015b, *ApJ*, 808, 139
- Sobral D., Kohn S. A., Best P. N., Smail I., Harrison C. M., Stott J., Calhau J., Matthee J., 2016, *MNRAS*, 457, 1739
- Song M. et al., 2014, *ApJ*, 791, 3
- Steidel C. C., Erb D. K., Shapley A. E., Pettini M., Reddy N., Bogosavljević M., Rudie G. C., Rakic O., 2010, *ApJ*, 717, 289
- Steidel C. C., Bogosavljević M., Shapley A. E., Kollmeier J. A., Reddy N. A., Erb D. K., Pettini M., 2011, *ApJ*, 736, 160
- Swinbank A. M., Sobral D., Smail I., Geach J. E., Best P. N., McCarthy I. G., Crain R. A., Theuns T., 2012, *MNRAS*, 426, 935
- Swinbank A. M. et al., 2015, *MNRAS*, 449, 1298
- Taniguchi Y. et al., 2015, *ApJ*, 809, L7
- Tasca L. A. M. et al., 2015, *A&A*, 581, A54
- Taylor M. B., 2005, in Shopbell P., Britton M., Ebert R., eds, *Astronomical Society of the Pacific Conference Series Vol. 347, Astronomical Data Analysis Software and Systems XIV*. Astron. Soc. Pac., San Francisco, p. 29

- Trainor R. F., Steidel C. C., Strom A. L., Rudie G. C., 2015, ApJ, 809, 89  
Umehata H. et al., 2015, ApJ, 815, L8  
Van Der Walt S., Colbert S. C., Varoquaux G., 2011, preprint ([arXiv:1102.1523](https://arxiv.org/abs/1102.1523))  
Venemans B. P. et al., 2007, A&A, 461, 823  
Verhamme A., Schaerer D., Atek H., Tapken C., 2008, A&A, 491, 89  
Wardlow J. L. et al., 2014, ApJ, 787, 9  
Wisotzki L. et al., 2015, preprint ([arXiv:1509.05143](https://arxiv.org/abs/1509.05143))  
Wold I. G. B., Barger A. J., Cowie L. L., 2014, ApJ, 783, 119  
Yang H., Malhotra S., Gronke M., Rhoads J. E., Jaskot A., Zheng Z., Dijkstra M., 2015, preprint ([arXiv:1506.02885](https://arxiv.org/abs/1506.02885))  
Zheng Z., Cen R., Trac H., Miralda-Escudé J., 2010, ApJ, 716, 574  
Zitrin A. et al., 2015, ApJ, 810, L12

This paper has been typeset from a  $\mathrm{T}_{\mathrm{E}}\mathrm{X}/\mathrm{L}^{\mathrm{A}}\mathrm{T}_{\mathrm{E}}\mathrm{X}$  file prepared by the author.